Lecture Notes in Networks and Systems 400

Amit Joshi · Mufti Mahmud ·
Roshan G. Ragel *Editors*

Information and Communication Technology for Competitive Strategies (ICTCS 2021)

ICT: Applications and Social Interfaces

Springer

Lecture Notes in Networks and Systems

Volume 400

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

Amit Joshi · Mufti Mahmud · Roshan G. Ragel
Editors

Information and Communication Technology for Competitive Strategies (ICTCS 2021)

ICT: Applications and Social Interfaces

Springer

Editors
Amit Joshi
Global Knowledge Research Foundation
Ahmedabad, India

Mufti Mahmud
Nottingham Trent University
Nottingham, UK

Roshan G. Ragel
University of Peradeniya
Kandy, Sri Lanka

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-19-0094-5 ISBN 978-981-19-0095-2 (eBook)
https://doi.org/10.1007/978-981-19-0095-2

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

Sixth International Conference on Information and Communication Technology for Competitive Strategies (ICTCS 2021) targets state-of-the-art as well as emerging topics pertaining to information and communication technologies (ICTs) and effective strategies for its implementation for engineering and intelligent applications.

The conference is anticipated to attract a large number of high-quality submissions, stimulate the cutting-edge research discussions among many academic pioneering researchers, scientists, industrial engineers, students from all around the world and provide a forum to researcher; propose new technologies, share their experiences and discuss future solutions for design infrastructure for ICT; provide a common platform for academic pioneering researchers, scientists, engineers and students to share their views and achievements; enrich technocrats and academicians by presenting their innovative and constructive ideas; and focus on innovative issues at international level by bringing together the experts from different countries.

The conference was held during December 17 and 18, 2021, physically at four points by Sheraton Jaipur, India, and digitally on Zoom organized by Global Knowledge Research Foundation.

Research submissions in various advanced technology areas were received, and after a rigorous peer review process with the help of programme committee members and external reviewer, 150 papers were accepted with an acceptance rate of 19%. All 150 papers of the conference are accommodated in 2 volumes; also, papers in the book comprise authors from 8 countries.

This event success was possible only with the help and support of our team and organizations. With immense pleasure and honour, we would like to express our sincere thanks to the authors for their remarkable contributions, all the technical program committee members for their time and expertise in reviewing the papers within a very tight schedule and the publisher Springer for their professional help.

We are overwhelmed by our distinguished scholars and appreciate them for accepting our invitation to join us through the virtual platform and deliver keynote speeches and technical session chairs for analyzing the research work presented by the researchers. Most importantly, we are also grateful to our local support team for

their hard work for the conference. This series has already been made a continuous series which will be hosted at different locations every year.

Ahmedabad, India — Amit Joshi
Nottingham, UK — Mufti Mahmud
Kandy, Sri Lanka — Roshan G. Ragel

Contents

Editors and Contributors

About the Editors

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

Dr. Mufti Mahmud received his Ph.D. degree in Information Engineering (specialized in neuroengineering) from the University of Padova, Italy, in 2011. Recipient of the Marie Curie postdoctoral fellowship, Dr. Mahmud has served at various positions in the industry and academia in India, Bangladesh, Italy, Belgium, and the UK during the last 17 years. Dr. Mahmud aims to leave behind a healthy, secure, and smart world to live in. As Expert of neuroengineering, computational intelligence, and data science, his research aims to build predictive, secure, and adaptive systems for personalized services to improve quality of life through advancement of healthcare access in low-resource settings. Senior Member of IEEE and ACM and Professional Member of BCS, Dr. Mahmud holds leadership roles at many technical committees, such as Vice Chair of the Intelligent System Application Technical Committee of IEEE Computational Intelligence Society, Member of the IEEE CIS Task Force on Intelligence Systems for Health, Co-chair of the IEEE

P2733 Standard on Intelligent systems design, with AI/ML, Member of the IEEE R8 Humanitarian Activities Subcommittee, and Project Liaison Officer of the IEEE UK and Ireland Special Interest Group on Humanitarian Technology. Dr. Mahmud serves as Associate Editor of the Cognitive Computation, IEEE Access, Brain Informatics, and Big Data Analytics journals. He also serves at many technical, program, and organization committees of high-rank conferences including the Local Organizing Chair of IEEE-WCCI2020; General Chair of BI2020; and Program Chair of IEEE-CICARE2020.

Dr. Roshan G. Ragel is Full Professor in Computer Engineering at the University of Peradeniya, Sri Lanka. He is Skilled University Lecturer and Researcher with a doctorate in computer science and engineering and a passion for teaching, research and administration. His research interest is in the areas of fault-tolerant and dependable computing, accelerated and high-performance computing, wearable computing, and synthesis of secure and reliable embedded processors. Prof. Ragel received his Ph.D. in Computer Science and Engineering from the University of New South Wales, Australia, in 2007 and his B.Sc. in Engineering from the University of Peradeniya, Sri Lanka, in 2001. He was appointed Senior Lecturer in Computer Engineering at the University of Peradeniya December 2007 and Full Professor since October 2017. Prof. Ragel has co-authored about 150 peer-reviewed articles in topics including micro-architectural support for reliability and security in embedded processors (SoC), Internet of Things (IoT), side-channel attacks and countermeasures, application-specific processor design, high-performance computing such as hardware/software acceleration of bioinformatics algorithms, wearable computing, format-preserving digitization and OCR of local language text, and completely automated public Turing test to tell computers and humans apart (CAPTCHAs) in the context of mobile applications. He has won several best paper awards in conference papers, the most prolific Sri Lankan Author Award from Elsevier Research in 2017, and the Presidential Award for scientific publication in Sri Lanka in 2018. He was Visiting Research Scholar at the UNSW, Sydney, a few times and UIUC, Illinois once. He is Senior Member of the IEEE and the IEEE Computer Society.

Contributors

Amina Adadi Moulay Ismail University, Meknes, Morocco

Ajay Agarwal KIET Group of Institutions, Ghaziabad, U.P., India

Pooja Agarwal PES University EC Campus, Bengaluru, India

Apeksha Aggarwal Jaypee Institute of Information Technology, Noida, U.P., India

Shivani Agrawal School of Electronics Engineering, Vellore Institute of Technology, Vellore, India

Ketut Agustini Universitas Pendidikan Ganesha, Singaraja, Bali, Indonesia

Shyam Akashe Department of Electronics and Communication Engineering, ITM University, Gwalior, M.P., India

K. R. Akshatha Department of Electronics and Communication Engineering, Centre of Excellence in Avionics and Navigation, Manipal, India;
Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India

Ketan Anand Department of Electronics and Instrumentation Engineering, M. S. Ramaiah Institute of Technology, Bengaluru, India

O. Anjali Rajadhani Institute of Science and Technology, Palakkad, Kerala, India

F. Antony Xavier Bronson Department of Biotechnology, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India

S. Anuradha National Institute of Technology Warangal, Warangal, Telangana, India

Chilupuri Anusha National Institute of Technology Warangal, Warangal, Telangana, India

P. Anvita KLE Technological University, Hubli, India

Ranjita Asati Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, India

S. Asha Vellore Institute of Technology Chennai Campus, Chennai, Tamil Nadu, India

Dipali Bansal GB Pant Okhla Campus 1, Delhi Skill & Entrepreneurship University, Delhi, India

Pratham Barve Birla Vishvakarma Mahavidyalaya Engineering College, Anand, Gujarat, India

Elena Basan Southern Federal University, Taganrog, Russian Federation

Amol Bhagat Department of Computer Science and Engineering, Prof Ram Meghe College of Engineering and Management, Badnera, Amravati, Maharashtra, India

Nandan Bhat Department of Computer Science Engineering, Manipal, India;
Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India

N. P. G. Bhavani Department of ECE, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India

Ankit Bhavsar GLS University, Ahmedabad, Gujarat, India

Kiran Bhowmick Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Pravin Kumar Bhoyar Symbiosis Institute of Management Studies, Symbiosis International (Deemed University), Pune, MH, India

Suneeta Bhudhihal SoECE, KLE Technological University, Hubli, Karnataka, India

Lyudmila V. Bunina MIREA—Russian Technological University, Moscow, Russia

K. Chaitra KLE Technological University, Hubli, India

Debangshu Chakraborty Agro-Computing Research Laboratory, Department of Computer Science, Ananda Chandra College, Jalpaiguri, India

Arti Chandani SIMS, Symbiosis International University, Pune, Maharashtra, India

Abhinandan Chandhok School of Electronics Engineering, Vellore Institute of Technology, Vellore, India

Aishwarya Chengalvala Department of Computer Science and Engineering, V R Siddhartha Engineering College, Vijayawada, India

Venkata Sairam Chennam Department of Computer Science and Engineering, Sasi Institute of Technology & Engineering, Tadepalligudem, Andhra Pradesh, India

Pradnya Chitrao Symbiosis Institute of Management Studies, A Constituent of Symbiosis International University, Pune, India

Rupali Chourey Department of CSE, SIRT, Bhopal, MP, India

Pranjal Chowdhury Institute of Engineering and Management, Kolkata, West Bengal, India

Bharti Chugh The NorthCap University, Gurugram, India

M. Deepika Department of Computer Science, Crescent University, Chennai, India

Makrand Deshpande Citicorp Services India Limited, Pune, India

Luh Joni Erawati Dewi Universitas Pendidikan Ganesha, Singaraja, Bali, Indonesia

Vu Thu Diep Hanoi University of Science and Technology, Hanoi, Vietnam

Rajiv Divekar Symbiosis Institute of Management Studies, Symbiosis International (Deemed University), Pune, MH, India

Vedant Dokania Birla Vishvakarma Mahavidyalaya Engineering College, Anand, Gujarat, India

Ruchi Dronawat Department of CSE, SIRT, Bhopal, MP, India

Pamela Vinitha Eric Professor, Presidency University, Bengaluru, India

Chenguttuvan Ezhilazhagan Dr. NGP Institute of Technology, Coimbatore, India

A. Ganesan Department of EEE, RRASE College of Engineering, Chennai, Tamil Nadu, India

C. K. M. Ganesan Vellore Institute of Technology Chennai Campus, Chennai, Tamil Nadu, India

I. Gede Aris Gunadi Universitas Pendidikan Ganesha, Singaraja, Bali, Indonesia

Jossy George CHRIST (Deemed to be University), Bengaluru, India

Nikhil George Department of Electronics and Telecommunication Engineering, Center for Imaging Technologies, M. S. Ramaiah Institute of Technology, Bengaluru, India

Indrajit Ghosh Agro-Computing Research Laboratory, Department of Computer Science, Ananda Chandra College, Jalpaiguri, India

Rajdeep Ghosh Institute of Engineering and Management, Kolkata, West Bengal, India

Leena Giri Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, India

Amit Grover Department of Electronics and Communication Engineering, Shaheed Bhagat Singh State University, Ferozepur, Punjab, India

Divya Sai Asanvitha Gundala Department of Computer Science and Engineering, V R Siddhartha Engineering College, Vijayawada, India

Anjani Gupta Dr.Akhilesh Das Gupta Institute of Technology and Management, Guru Gobind Singh Indraprastha University, New Delhi, India

Kriti Gupta Arya College of Engineering and IT Jaipur, Jaipur, India

M. Guruprasath Computer Science and Engineering, Vel Tech Rangarajan Dr, Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India

Aayush Halgekar Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Rajeswary Hari Department of Biotechnology, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India

Nallamothu Lakshmi Hemanth Department of Computer Science and Engineering, V R Siddhartha Engineering College, Vijayawada, India

U. Hemavathi Department of Computer Science and Egineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India

M. Himavanth Sai Computer Science and Engineering, Vel Tech Rangarajan Dr, Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India

Shraddha Hiremath SoECE, KLE Technological University, Hubli, Karnataka, India

Muthyala Hithasri Kakatiya Institute of Technology and Science, Warangal, Telangana, India

Shivaraj Hublikar K.L.E. Technological University, Hubli, India

Phan Duy Hung FPT University, Hanoi, Vietnam

Gede Indrawan Universitas Pendidikan Ganesha, Singaraja, Bali, Indonesia

C. A. Irfana Parveen Sahrdaya College of Engineering and Technology, Thrissur, Kerala, India

Md. Ismail Jabiullah Daffodil International University, Dhaka, Bangladesh, India

Nalini C. Iyer K.L.E. Technological University, Hubli, India

Arpit A. Jain GLS University, Gujarat, India

Dhyanendra Jain Dr.Akhilesh Das Gupta Institute of Technology and Management, Guru Gobind Singh Indraprastha University, New Delhi, India

U. Jayalatsumi Departmentt of ECE/Civil Engineering, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India

Arun Kumar Jhapate Department of CSE, SIRT, Bhopal, MP, India

Dennis George Johnson Department of Computer Science Engineering, Manipal, India;
Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India

Solley Joseph CHRIST (Deemed to be University), Bengaluru, India;
Carmel College of Arts, Science and Commerce for Women, Goa, India

Nandita Kadam K. J. Somaiya College of Engineering, Vidyavihar, Mumbai, India

Sethu Sandeep Kadiyala Department of Computer Science and Engineering, V R Siddhartha Engineering College, Vijayawada, India

A. Kalaivani Department of CSE, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India

Pille Kamakshi Kakatiya Institute of Technology and Science, Warangal, India

R. Kanagavalli Information Science and Engineering, The Oxford College of Engineering, Bengaluru, Karnataka, India

Krishna Kant Smt. Chandaben Mohanbhai Patel Institute of Computer Applications (CMPICA), Charusat University, Changa, Anand, Gujarat, India

D. Karthika Department of Computer Science, SDNB Vaishnav College for Women, Chennai, India

A. K. Karunakar Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India;
Department of Data Science and Computer Applications, Manipal, India

Jenny Kasudiya GLS University, Ahmedabad, Gujarat, India

T. Kavitha Departmentt of ECE/Civil Engineering, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India

I. Ketut Paramarta Universitas Pendidikan Ganesha, Singaraja, Bali, Indonesia

Rustam F. Khalabiya MIREA—Russian Technological University, Moscow, Russia

Dhruvi Khankhoje Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Roshni S. Khedgaonkar Department of Computer Technology, Yeshwatrao Chavan College of Engineering, Nagpur, Maharashtra, India

Ishaan Khetan Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Bhagyashree Kinnal K.L.E. Technological University, Hubli, India

Kishore O Aditya Engineering College, Surampalem, India

Shweta Kore JSS College Dharwad, Dharwad, Karnataka, India

Pranali Kosamkar MIT World Peace University, Pune, India

Bhavak Kotak Parul University, Vadodara, India

Ben Kotze Central University Technology, Bloemfontein, Free State, South Africa

G. Kranthi Kumar Department of Computer Science and Engineering, V R Siddhartha Engineering College, Vijayawada, India

R. Krishnakumar School of Electrical Engineering, Vels Institute of Science, Technology and Advanced Studies, Chennai, India

Ashwin R. Kubasadgoudar KLE Technological University, Hubli, India

Abhishek Kuber MIT World Peace University, Pune, India

Sravana Sandhya Kuchu Department of Computer Science and Engineering, Sasi Institute of Technology & Engineering, Tadepalligudem, Andhra Pradesh, India

Soham Kulthe MIT World Peace University, Pune, India

Aarti Kumar Rabindra Nath Tagore University, Bhopal, Madhya Pradesh, India

Atul Kumar Parul University, Vadodara, India

Thota Mahesh Kumar Kakatiya Institute of Technology and Science, Warangal, India

Meet Kumari Department of Electronics and Communication Engineering, Chandigarh University, Mohali, Punjab, India

Rangith Baby Kuriakose Central University Technology, Bloemfontein, Free State, South Africa

L. S. N. J. Manjusha P Aditya Engineering College, Surampalem, India

Mohamed Lahmer Moulay Ismail University, Meknes, Morocco

Manjunath Lakkannavar MSRIT, Bengaluru, India

Karuppiah Lakshmi Prabha PSG Polytechnic College, Coimbatore, India

B. Lakshmi KLE Technological University, Hubli, India

Ntebaleng Junia Lemphane Central University Technology, Bloemfontein, Free State, South Africa

Krishn Limbachiya Department of Instrumentation and Control Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

S. Maheswari School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India

Urmila Mahor Rabindra Nath Tagore University, Bhopal, Madhya Pradesh, India

Nitin Malik The NorthCap University, Gurugram, India

S. M. Manasa Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, India

Venkatesh Mane K.L.E. Technological University, Hubli, India

Nikhil S. Mangrulkar Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India

S. H. Manjula Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, India

Ann C. V. Medona Department of Artificial Intelligence and Data Science, Saveetha Engineering College, Chennai, Tamil Nadu, India

Krushil Mehta Birla Vishvakarma Mahavidyalaya Engineering College, Anand, Gujarat, India

Vishant Mehta K. J. Somaiya College of Engineering, Vidyavihar, Mumbai, India

Sunanda Mendiratta Department of Electronics Engineering, JCBoseUST, Faridabad, India

Jasmine Minj Department of Computer Science and Engineering, VEC, Lakhanpur, CG, India

Kavita Mittal Jagannath University NCR, Bahadurgarh, Haryana, India

F. Monica Rexy Vellore Institute of Technology, Vellore, Tamil Nadu, India

Shreyas More K. J. Somaiya College of Engineering, Vidyavihar, Mumbai, India

Md. Mynul Islam Daffodil International University, Dhaka, Bangladesh, India

A. Nagarajan Department of Computer Applications, Alagappa University, Karaikudi, India

Siddharth Nair Vellore Institute of Technology Chennai Campus, Chennai, Tamil Nadu, India

Le Hoang Nam FPT University, Hanoi, Vietnam

Samia Nasiri Moulay Ismail University, Meknes, Morocco

Shravani Naware MIT World Peace University, Pune, India

B. Neeraja SIMS, Symbiosis International University, Pune, MH, India

P. Nikita KLE Technological University, Hubli, India

Sravanthi Nittala PES University EC Campus, Bengaluru, India

Elena V. Nurmatova MIREA—Russian Technological University, Moscow, Russia

Umang Pachaury Parul University, Vadodara, India

M. K. Padmapriya Research Scholar, New Horizon College of Engineering, Bengaluru, India

G. Padmashree Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India

Krisha Panchamia K. J. Somaiya College of Engineering, Vidyavihar, Mumbai, India

A. Pandian SRM Institute of Science and Technology, Kattankulathur, India

Kokila Paramanandam Computer Science & Engineering, PES University, Bengaluru, Karnataka, India

Madhuri Parekh Computer Engineering Department, Ph.d scholar Marwadi University, Rajkot, India

Shahee Parveen Department of Computer Science and Engineering, VEC, Lakhanpur, CG, India

Dev Patel Birla Vishvakarma Mahavidyalaya Engineering College, Anand, Gujarat, India

Preet Patel Parul University, Vadodara, India

Pooja Patre Department of Computer Science and Engineering, VEC, Lakhanpur, CG, India

Vishal Pattanashetty SoECE, KLE Technological University, Hubli, Karnataka, India

Vishal B. Pattanashetty K.L.E. Technological University, Hubli, India

Sourav Paul Institute of Engineering and Management, Kolkata, West Bengal, India

A. Peter Soosai Anandarai Computer Science and Engineering, Vel Tech Rangarajan Dr, Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India

Preeti S. Pillai K.L.E. Technological University, Hubli, India

Kamakshi Pille Kakatiya Institute of Technology and Science, Warangal, Telangana, India

P. Prabhavathy Vellore Institute of Technology, Vellore, Tamil Nadu, India

Om Prakash SICSR, Symbiosis International University, Pune, Maharashtra, India

G. Priya Vellore Institute of Technology, Vellore, Tamil Nadu, India

Nikita Proshkin Southern Federal University, Taganrog, Russian Federation

M. M. Raghuwanshi Department of Data Science, G.H. Raisoni College of Engineering and Management, Pune, India

Mukesh M. Raghuwanshi Symbiosis Institute of Technology, Symbiosis International (Deemed University), Nagpur, Maharashtra, India

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

V. Ramu Jawaharlal Nehru Technological University Hyderabad, Hyderabad, India

Ashish Rao Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Sanskruti Raut MIT World Peace University, Pune, India

V. G. Renumol Cochin University of Science and Technology, Cochin, Kerala, India

Yassine Rhazali Moulay Ismail University, Meknes, Morocco

Rahul Rudra Institute of Engineering and Management, Kolkata, West Bengal, India

S.S. Sunaina M Aditya Engineering College, Surampalem, India

S. Sabarish Vellore Institute of Technology Chennai Campus, Chennai, Tamil Nadu, India

Mohd Safiya Kakatiya Institute of Technology and Science, Warangal, India

S. Sahithya Rajalakshmi Engineering College, Chennai, India

Pawankumar Saini Symbiosis International University, Pune, India

S. Sam blesswin Rajalakshmi Engineering College, Chennai, India

Neha A. Samsir GLS University, Gujarat, India

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

P. Sasikumar School of Electronics Engineering, Vellore Institute of Technology, Vellore, India

P. Satish Kumar University College of Engineering, Osmania University, Hyderabad, India

B. Satish Shenoy Department of Aeronautical and Automobile Engineering, Centre of Excellence in Avionics and Navigation, Manipal, India;
Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India

Mit Savani Department of Instrumentation and Control Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Sanyam Savla K. J. Somaiya College of Engineering, Vidyavihar, Mumbai, India

Minal Saxena Department of EC, SIRT, Bhopal, MP, India

K. Senthil Kumar Department of ECE, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India

T. Senthil Murugan Kakatiya Institute of Technology and Science, Warangal, India

Kathiresan Senthil Department of Electronics and Instrumentation Engineering, M. S. Ramaiah Institute of Technology, Bengaluru, India

Mayur M. Sevak Birla Vishvakarma Mahavidyalaya Engineering College, Anand, Gujarat, India

Sk. Shafiya Department of ECE, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India;
Nalla Narasimha Reddy Group of Institutions, Hyderabad, India

D. B. Shah G.H Patel P G Department of Computer Science and Technology, Sardar Patel University, Anand, Gujarat, India

Parth Shah Parul University, Vadodara, India

K. Shamshuddin SoECE, KLE Technological University, Hubli, Karnataka, India

Sahana Shanbhag PES University EC Campus, Bengaluru, India

Ankit Sharma Department of Instrumentation and Control Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Anupam Kumar Sharma Dr.Akhilesh Das Gupta Institute of Technology and Management, Guru Gobind Singh Indraprastha University, New Delhi, India

Reecha Sharma Department of Electronics and Communication Engineering, Punjabi University, Patiala, Punjab, India

Roohi Sharma CHRIST (Deemed to Be University), Bengaluru, India

Anu Sheetal Department of Engineering and Technology, Guru Nanak Dev University, Gurdaspur, Punjab, India

N. S. V. Shet NIT-Karnataka, Mangaluru, India

Shraddha KLE Technological University, Hubli, India

Shreesha Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, India

Ankit Shrivastava SCIT, Manipal University Jaipur, Jaipur, Rajasthan, India

Vishal Shrivastava Arya College of Engineering and IT Jaipur, Jaipur, India

Aditi Shukla SCIT, Manipal University Jaipur, Jaipur, Rajasthan, India

Madhu Shukla Computer Engineering Department, Associate Professor & Head Computer Engineering-AI, Marwadi University, Rajkot, India

M. Shyamala Devi Computer Science and Engineering, Vel Tech Rangarajan Dr, Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India

Amala Siby CHRIST (Deemed to Be University), Bengaluru, India

Kavita R. Singh Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India

Prashant Singh Dr.Akhilesh Das Gupta Institute of Technology and Management, Guru Gobind Singh Indraprastha University, New Delhi, India

Akula V. S. Siva Rama Rao Department of Computer Science and Engineering, Sasi Institute of Technology & Engineering, Tadepalligudem, Andhra Pradesh, India

D. Sorna Shanthi Rajalakshmi Engineering College, Chennai, India

R. Sreevarshini Rajalakshmi Engineering College, Chennai, India

Aditya Srinivas SIMS, Symbiosis International University, Pune, MH, India

G. N. Srinivas Jawaharlal Nehru Technological University Hyderabad, Hyderabad, India

Devesh Kumar Srivastava SCIT, Manipal University Jaipur, Jaipur, Rajasthan, India

N. Srividya SIMS, Symbiosis International University, Pune, MH, India

Irina V. Stepanova MIREA—Russian Technological University, Moscow, Russia

K. Sujatha Department of EEE, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India

Farzana Sultana Daffodil International University, Dhaka, Bangladesh, India

R. Sunder Sahrdaya College of Engineering and Technology, Thrissur, Kerala, India

B. Suneeta KLE Technological University, Hubli, India

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

P. Swathi National Institute of Technology, Karnataka, Mangaluru, India

Viswanath Talasila Department of Electronics and Telecommunication Engineering, Center for Imaging Technologies, M. S. Ramaiah Institute of Technology, Bengaluru, India

Senthil Murugan Tamilarasan Kakatiya Institute of Technology and Science, Warangal, Telangana, India

Md. Tanbir Hasan Daffodil International University, Dhaka, Bangladesh, India

Vanita Tank MIT World Peace University, Pune, India

R. Tanuja Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, India

Daivakrupa Thota Department of Computer Science and Engineering, Sasi Institute of Technology & Engineering, Tadepalligudem, Andhra Pradesh, India

Neelam Turk Department of Electronics Engineering, JCBoseUST, Faridabad, India

Dhrumil Vadera Department of Instrumentation and Control Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

J. Vasanth Wason Department of Computer Applications, Alagappa University, Karaikudi, India

R. Vasanth SRM Institute of Science and Technology, Kattankulathur, India

Prashant Vats Dr.Akhilesh Das Gupta Institute of Technology and Management, Guru Gobind Singh Indraprastha University, New Delhi, India

Teena Vats Guru Gobind Singh Indraprastha University, Dwarka, Delhi, India

M. Venkatesan National Institute of Technology, Karnataka, Mangaluru, India

K. R. Venugopal Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru, India

Priya Vijay Rajalakshmi Engineering College, Chennai, India

Bui Trong Vinh Hanoi Procuratorate University, Hanoi, Vietnam

V. V. Vinu Varghese Cochin University of Science and Technology, Cochin, Kerala, India

P. Vishal KLE Technological University, Hubli, India

R. Vishnu PES University EC Campus, Bengaluru, India

Smita Wagholikar SIT, Symbiosis International University, Pune, Maharashtra, India

Haritha Yantrapragada Department of Computer Science and Engineering, Sasi Institute of Technology & Engineering, Tadepalligudem, Andhra Pradesh, India

Evgeniy I. Zaytsev MIREA—Russian Technological University, Moscow, Russia

Two-terminal Reliability Assessments of a Complex System

Ajay Agarwal and Apeksha Aggarwal

Abstract Assessment of the reliability of a system from its basic elements is one of the most important aspects of reliability analysis. To determine overall reliability easily, the given system must have all components organized either in series or in parallel. But if that is not the case, calculating the overall reliability of a system becomes challenging. In a complex system, one can group the components of the system neither in series nor in parallel. Therefore, the determination of the reliability of complex systems becomes more complicated. There are two ways for computing network reliability: exact and approximate. The approaches discussed in this paper compute exact network reliability.

Keywords Assessment · Complex system · Reliability

1 Introduction

Nowadays, one person communicates with others by using various media either through phone calls or through email or through any social media platform. Such medium is connected via a communication network whether it is a telecommunication network/Internet. These networks are consisting of nodes (system components) and links that connect the nodes. These links among nodes consist of both hardware and software components which make communication among nodes possible in such networks. Therefore, it becomes significant to learn the reliability of such a network so that communication [1] will not get stopped in the event of the failure of one/more components.

A. Agarwal (✉)
KIET Group of Institutions, Ghaziabad, U.P., India
e-mail: ajay.agarwal@kiet.edu

A. Aggarwal
Jaypee Institute of Information Technology, Noida, U.P., India
e-mail: apeksha.aggarwal@mail.jiit.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_1

In reality, such networks are represented in which nodes may be connected in series or in parallel or in a combination of both. If these nodes are connected either in series or in parallel, evaluation of reliability [2] can be done easily using the mathematical formulas available. But, this is often the case that a network model [3] of a system has components linked neither in series nor in parallel. As a result, such systems are hard to analyze. In that case, we have to explore techniques that will help to convert such complex networks into simple series or parallel components in such a way so that overall reliability cannot be compromised. Four of the methods used for exploration are explained in detail in Sect. 3. In Sect. 2, network graph model is defined. Evaluation approaches are discussed in Sect. 3. Concluding remarks are made in Sect. 4.

2 Network Graph Model

Communication networks [4] are usually modeled by graphs [1, 5]. A graph G is defined as $G = (V, E)$ where V is a set of nodes (network elements) and E is a set of links (unordered pairs of distinct nodes) [6] with the probabilities of node/edge being operational. In subsequent sections, our focus will be on the point that the communication between two designated nodes a source S and destination D must take place in spite of some node/link failure(s). In this paper, we will assume that all nodes are completely reliable. So, there will be a need only to explore link failures methods.

Network Reliability: In this paper, reliability between source node S and the destination node is considered. It is termed as the reliability of two-terminal network [5]. The reliability of such a network is defined as the probability that these two nodes are connected by some path of the reliable and operational links. Analytical methods to evaluate two-terminal reliability are discussed in the following sections.

3 Evaluation Methods

Series and parallel networks are defined below:

3.1 Series Network

The components in a network are said to be connected in series if they all must work/operational for communication to succeed between source S and destination D. If at least one component gets failed, the communication between S and D gets failed. Consider a network (see Fig. 1) having two components a and b connected in series. Further, let

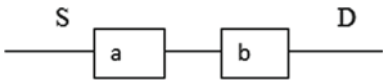


Fig. 1 Components *a* and *b* connected in series between *S* and *D*

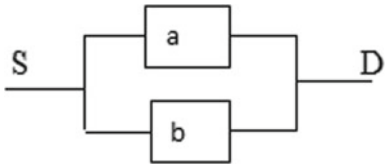


Fig. 2 Components *a* and *b* connected in parallel between *S* and *D*

– R_a = probability that component *a* is working/operational.

As success and failure are mutually exclusive, we have

– $Q_a = 1 - R_a$ = probability that the component *a* is not working/operational.
– R_b = probability that component *b* is working/operational and $Q_b = 1 - R_b$.

Probability of communication success or reliability R_s is that both components *a* and *b* must be working/operational, given by Eq. (1) as follows:

$$R_s = R_a \cdot R_b \tag{1}$$

If there exists *n* components/nodes, viz., 1, 2, 3, … *n* in series, then Eq. (1) can be generalized to give $R_s = R_1 \cdot R_2 \cdot R_3 \ldots R_n$ as per Eq. (2):

$$Q_s = 1 - R_s \tag{2}$$

Note: As each component has reliability less than 1, the product (of reliability of each component) will result in a value of less than 1. Reliability R_s is inversely proportional series components available.

3.2 Parallel Network

The nodes/components are said to be in parallel (see Fig. 2) if communication from source to destination fails if all parallel components/nodes get failed. A parallel system with two components *a* and *b* is shown in Fig. 2. Reliability of communication success between *S* and *D* will naturally be if one component (either *A* or *B* or both) need be working. The reliability equation for parallel network is given in Eq. (3) as follows:

$$R_s = 1 - Q_a \cdot Q_b = 1 - (1 - R_a)(1 - R_b) = R_a + R_b - R_a R_b \tag{3}$$

Or for an n component system, described mathematically by Eq. (4) as follows:

$$R_s = 1 - (Q_1 Q_2 \cdot Q_3 \ldots Q_n) = 1 - Q_s \tag{4}$$

where $Q_s = Q_1 Q_2 \cdot Q_3 \ldots Q_n$. *Note: Reliability R_s is directly proportional to parallel components available* ***(opposite to series)***.

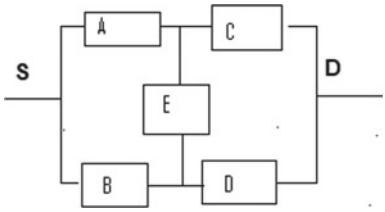
3.3 Complex (Non-series/Parallel) System

Generally, it is difficult to find series components and parallel components in the given complex or meshed network. Figure 3 is an example of such network. In Fig. 3, the network can be divided neither in series nor in parallel networks. Therefore, one cannot use Eqs. (1), (2), (3), and (4). Now, we will explore technique, such that the entire network can be grouped into series/parallel configurations so that we can evaluate communication success/reliability analytically/mathematically with the help of these formulas. In this paper, it is assumed that the reliability of each node/component is known a priori. Using grouping components, several techniques exist so that reliability of a two-terminal network can be determined.

State (Event) Space Method: In this method, representing all possible states of the network is called as state space of the network [7]. The system state describes the states of its components. All possible $2|E|$ (where $|E|$ implies the number of components present in the system) system states (or events) if each of the components (edges) in E assumes two states operational and non-operational. The communication success/reliability of the two-terminal network is the probability of the union of all mutually exclusive events that provide communication success in taking the message from the input terminal to the output terminal. On the other hand, failure of two-terminal network communication is the union of all mutually exclusive events that do not provide a communication success/failure from source to destination node.

Reliability of the Bridge Network: In Fig. 3, there are five components (A, B, C, D, E), having two states (success as 1, failure as 0). Therefore, there will be $2^5 = 32$ events. Out of 32 events, only the following 16 events provide communication success between S and D. Overall network reliability is the probability of the union of all these events S and D. Using Eqs. (1) and (2), reliability of the bridge network is given as per Eq. (5). Let reliability of each node be represented by 1 and 0 of all the five nodes R_A, R_B, R_C, R_D, and R_E. Then, communication success will be achieved if R_A, R_B, R_C, R_D, R_E.

Fig. 3 Bridge network



$$
\begin{aligned}
R_s = {} & (1-R_A)R_B(1-R_C)R_D(1-R_E) + (1-R_A)R_B(1-R_C)R_DR_E \\
& + (1-R_A)R_BR_C(1-R_D)R_E + (1-R_A)R_BR_CR_D(1-R_E) \\
& + (1-R_A)R_BR_CR_DR_E + R_A(1-R_B)(1-R_C)R_DR_E \\
& + R_A(1-R_B)R_C(1-R_D)(1-R_E) + R_A(1-R_B)R_C(1-R_D)R_E \\
& + R_A(1-R_B)R_CR_D(1-R_E) + R_A(1-R_B)R_CR_DR_E \\
& + R_AR_B(1-R_C)R_D(1-R_E) + R_AR_B(1-R_C)R_DR_E \\
& + R_AR_BR_C(1-R_D)(1-R_E) + R_AR_BR_C(1-R_D)R_E \\
& + R_AR_BR_CR_D(1-R_E) + R_AR_BR_CR_DR_E
\end{aligned} \tag{5}
$$

$$
\begin{aligned}
R_s &= 2(0.9^2 0.1^3) + 8(0.9^3 0.1^2) + 5(0.9^4 0.1^1) + 1(0.9^5) \\
&= 2 * 0.00081 + 8 * 0.00729 + 5 * 0.06561 + 0.59049 \\
&= 0.00162 + 0.05832 + 0.32805 + 0.59049 = \mathbf{0.97848}
\end{aligned}
$$

Decomposition Method: This method utilizes the reduction of the system by imposing conditions upon the branch(es) those destroying the series–parallel structure [8]. At the end of sequential reductions, complex connected system will be decomposed into several series–parallel connected subsystems.

Reliability of the Bridge Network: If a condition is applied to unit E, bridge network can be broken into two sub-networks. In one sub-network, link E is considered working/reliable (short circuit with 0 reliability), and in the other sub-network, E is considered failed (open circuit). Subdivision (see Fig. 4) and associated equation (Eq. 6) are as follows:

$$Rs = R_{\mathrm{I}}R_E + R_{\mathrm{II}}Q_E \tag{6}$$

where R_{I} and R_{II} are reliabilities of Subsystems I and II, respectively, as given in Fig. 4. If each component has the same reliability = 0.9, then $R_s = 0.88209 + 0.09639 = \mathbf{0.97848}$ (same value as obtained in the previous method).

Cut-set Method: Before giving the details of the method [9, 10], we will define several terms. Cut-set: If a set of links in a two-terminal network failed, communication success will fail. Minimal cut-set (MCS): If the set of links in a two-terminal

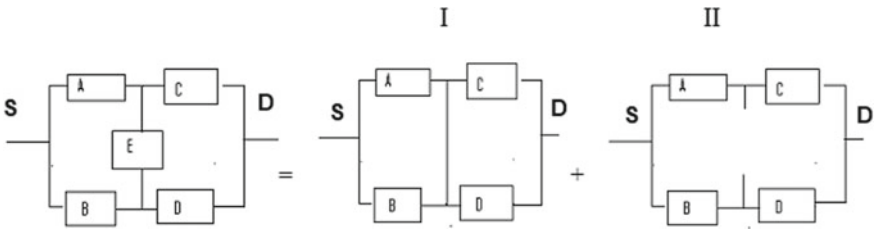


Fig. 4 Subdivision of bridge network

network failed, communication success will fail, but if any one link in MCS is not failed, communication success will be achieved. The following conclusions can be drawn from the definition of MCS.

1. As all the components of a MCS are necessary to fail for system failure, these components can be understood to be connected *parallel* among themselves.
2. There are several MCSs of a given a complex two-terminal network. Since the failure of one MCS is sufficient for system failure, MCSs of the system are *serially* connected among themselves.

According to these two conclusions, a complex network can be represented by the serially connected MCSs of the network. The resulting equivalent comprising serially connected MCSs is a series–parallel structure. However, there is an important difference from real series–parallel structure. Several components may appear two or more times in the structure since they may be included by two or more MCSs. Therefore, failure probabilities of MCSs comprising common elements are not independent of each other. This brings the usage of "union" concept in reliability evaluation. Network reliability can be expressed in terms of MCS reliabilities as,

$$R_s = \mathrm{REL}(C_1) \cdot [\mathrm{REL}(C_2)|\mathrm{REL}(C_1)] \cdot [\mathrm{REL}(C_3)|\{\mathrm{REL}(C_1), \mathrm{REL}(C_2)\}] \ldots [\mathrm{REL}(C_n)|\{\mathrm{REL}(C_1), \mathrm{REL}(C_2), \ldots, \mathrm{REL}(C_{n-1})\}] \tag{7}$$

This expression in Eq. (7) requires so many calculations. Instead, expression of failure probability is preferred, which is given in Eq. (8) as follows:

$$\begin{aligned} Q_s &= Q[C_1 U C_2 U C_3 U \ldots U C_n] \\ &= \sum_{i=1}^{n} Q(Ci) - \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} Q(Ci) \cap Q(Cj) \\ &\quad + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} Q(Ci \cap Cj \cap Ck) \end{aligned} \tag{8}$$

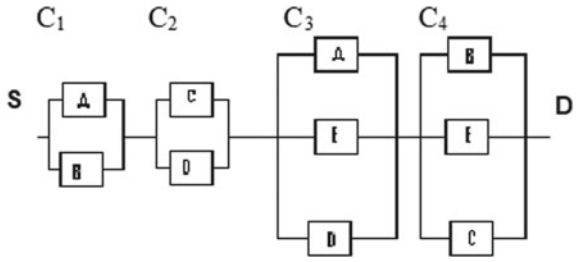


Fig. 5 Bridge network equivalent in terms of MCSs

Note that the failure probability of an intersection of two MCSs is determined by taking the product of the failure probabilities of the links included by those MCSs.

Three and higher-order intersections of MCSs can similarly be determined. However, failure probabilities of MCS intersections decrease as the order of intersection increases. Therefore, the lowest order failure products and one higher-order product are assumed to be enough for a reasonable accuracy.

Reliability of the Bridge Network: MCSs of the bridge network are *AB*, *CD*, *AED*, and *BEC* (see Fig. 5). Corresponding equation is given in Eq. (9).

$$\begin{aligned} Q_s = \text{Failure of } [C_1 U C_2 U C_3 U C_4] &= Q_A Q_B + Q_C Q_D + Q_A Q_E Q_D \\ &+ Q_B Q_E Q_C - [Q_A Q_B Q_C Q_D + Q_A Q_B Q_D Q_E + Q_A Q_B Q_C Q_E \\ &+ Q_A Q_C Q_D Q_E + Q_B Q_C Q_D Q_E + Q_A Q_B Q_C Q_D Q_E] \\ &+ 4 * Q_A Q_B Q_C Q_D Q_E - Q_A Q_B Q_C Q_D Q_E \end{aligned} \tag{9}$$

If the failure probability of all units in the bridge network is $Q = 0.1$; $Q_s = 2 * 0.1^2 + 2 * 0.1^3 - [5 * 0.1^4 + 0.1^5] + [4 * 0.1^5] - 0.1^5 = 0.022 - 0.0005 + 0.00002 = 0.02150 + 0.00002 = 0.02152 R_s = 0.97848$ (same as obtained in previous methods). The lowest order failure product is two (*AB* or *CD*). Therefore, failure products up to third order are enough for a reasonable accuracy. $Q_s \approx Q_A Q_B + Q_C Q_D + Q_A Q_E Q_D + Q_B Q_E Q_C$. Thus, $Q_s = 0.1^2 + 0.1^2 + 0.1^3 + 0.1^3 = 0.02 + 0.002 = 0.022$; $R_s = 0.978$.

Path-tracing Method (Path Set or Tie Set): Path-tracing method [11, 12] is complementary to the cut-set method. Terms need to be defined first:

Tie Set: A set of links connected in series. The presence of all the links in any tie set connects nodes *S* and *D*. The success of the system to carry messages from source *S* to output *D* depends on the success of each link of any tie set.
MTS: Set of tie sets where the removal of any of its link is no longer considered a tie set. This implies that all links of an MTS must succeed to cause the communication success.

It is to be noted that a single tie set is enough for system operation. If there exist multiple tie sets, then they must be connected in parallel among themselves.

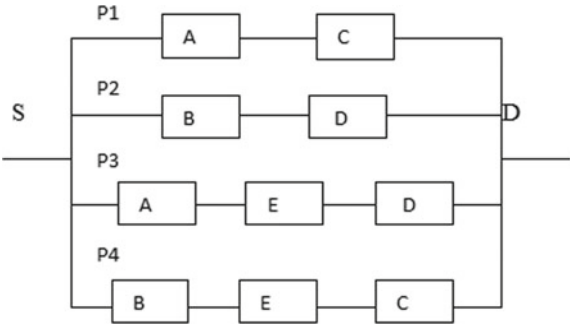


Fig. 6 Tie set equivalent of the bridge network

Reliability of the Bridge Network: Minimal paths of the bridge network were *AC*, *BD*, *AED*, and *BEC*. The tie set equivalent of the system is given in Fig. 6.

Note that the evaluations are the similar ones those applied in minimal cut-set method. However, here is an important difficulty. Since the reliabilities of higher-order intersections are *not small* enough to be neglected, all higher-order intersection reliabilities need to be calculated. This computational inefficiency together with the lack of information about failure modes of the system is the most important two drawbacks of the method.

The communication success/reliability of the bridge network between nodes *S* and *D* is evaluated as per Eq. (10).

$$\begin{aligned} R_s = \text{Reliability of Set } [P_1 U P_2 U P_3 U P_4] &= R_A R_B + R_C R_D + R_A R_E R_D \\ &+ R_B R_E R_C - [R_A R_B R_C R_D + R_A R_B R_D R_E + R_A R_B R_C R_E \\ &+ R_A R_C R_D R_E + R_B R_C R_D R_E + R_A R_B R_C R_D R_E] \\ &+ 4 * R_A R_B R_C R_D R_E - R_A R_B R_C R_D R_E \end{aligned} \tag{10}$$

If the reliability of all units in the bridge network is $R = 0.9 R_s = 2[0.9^2 + 0.9;^3] - [5 \times 0.9^4 + 0.9^5] + [4 \times 0.9^5] - 0.9^5 = 2[1.5390] - [5 \times 0.9^4] + 2 \times 0.9^5 = 3.07800 - 3.28050 + 1.18098 = 4.25898 - 3.28050$ $R_s =$ **0.97848** (same as obtained in previous methods).

4 Conclusion

Terminal reliability assessment techniques of a complex system are discussed. To measure the accurate evaluation of communication success between given two nodes, state space method is used. In this method, every different formulation of the network is considered. It is conceptually general and flexible. But, it is time-consuming as state space is exponential. On the other hand, the decomposition method is consid-

ered useful for a small network and, therefore, is used to evaluate the reliability of communication success in many applications. But, the process is not adequate for computer-aided reliability evaluation of large-scale complex systems.

If not the entire network is subject of interest, then the minimal cut-set method is a useful tool. This tool significantly reduces computation time. As opposed to the simulation method [4], analytical methods have the following advantages: (i) In case of using simulation, the result depends on the number of simulations. (ii) In systems with components in which the reliability does not change with time cannot be simulated. (iii) In case of analytical method, a mathematical expression is obtained to determine the reliability of the system.

The simulation method is used in case of a large network with many components where it becomes hard to determine reliability of the communication success between given two nodes analytically even with the use of computers.

References

1. Ram M, Pham H (2020) Advances in reliability analysis and its applications. Springer International Publishing
2. Anders G (1990) Probability concepts in electric power systems. Wiley, New York
3. Billinton R, Allan R (1992) Reliability evaluation of engineering systems: concepts and techniques, 2nd edn. Plenum Press, New York
4. Zarghami SA, Gunawan I, Schultmann F (2020) Exact reliability evaluation of infrastructure networks using graph theory. Qual Reliab Eng Int 36(2):498–510
5. Chaturvedi SK, Khanna G, Soh S (2018) Reliability evaluation of time evolving delay tolerant networks based on sum-of-disjoint products. Reliab Eng Syst Saf 171:136–151
6. Berge C (1962) The theory of graphs. Methuen, London
7. Nikolovski S, Sljivac D, Stefic B (1999) State enumeration approach in reliability assessment of the Eastern Croatia bulk power system. Electric Power Eng
8. Lucet C, Manouvrier J. Exact methods to compute network reliability. http://citeseer.ist.psu.edu/432828.html
9. Shooman ML (2002) Reliability of computer systems and networks: fault tolerance, analysis, and design, chap 6. Wiley. ISBN 0-471-29342-3
10. Medhi D (2007) Network reliability and fault-tolerance. In: Wiley encyclopedia of computer science and engineering
11. Médard M (2003) Network reliability and fault tolerance. In: Encyclopedia of telecommunications. Wiley Online Library
12. Gupta AK. Reliability, maintenance and safety engineering. University Science Press, Golden House, Daryaganj, New Delhi, p 113

AI-Based Interactive Agent for Health Care Using NLP and Deep Learning

U. Hemavathi and Ann C. V. Medona

Abstract Artificial intelligence (AI)-based interactive agent plays a vital role in health care. In this paper, we have implemented AI-based Interactive agent using natural language processing (NLP) and deep learning (DL) which deals with resolving simple queries and provide health care services. This work helps to build a system for providing better health care service for user and increase user interaction. This work builds a chatbot using DL and NLP pipeline. We conclude that chatbots build with NLP and neural network are more efficient than human resource and predefined chatbots that are not efficient. It is necessary to discuss potential research directions that can improve capabilities.

Keywords Deep learning · Artificial intelligence · Natural language processing · Health care

1 Introduction

AI chatbot is intelligent software that can provide efficient communication and perform actions similar to human beings [1]. It yields various services such as marketing on social networks and effective customer interaction. In recent periods, chatbots give answers to the client's questionnaires for any particular domain where it is used. It can accept speech and text inputs for its operation. Chatbots are mainly used for information acquisition. These AI embedded agents use natural language processing techniques to answer the user questions. Predefined knowledge base helps develop a response to the query [2, 3]. There is a need for such mechanisms to

U. Hemavathi (✉)
Department of Computer Science and Egineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India
e-mail: uhemav@gmail.com

A. C. V. Medona
Department of Artificial Intelligence and Data Science, Saveetha Engineering College, Chennai, Tamil Nadu, India
e-mail: medona@saveetha.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_2

determine the response to query from available data. Hence, to solve the problem, feed forward neural network-based mechanism is proposed which can analyze the response and reply the user [4, 5].

1.1 Scope of the Work

AI chatbots are time efficient and cost effective which provides efficient services for 24*7. This work proposes to implement AI-based interactive agent for health care using NLP and deep learning which deal with resolving simple queries and provide health care service. Deep learning techniques use artificial intelligence (AI) that can make the computer to learn from the data instead of being explicitly programmed. Training data are processed, and features are extracted by ML algorithm which is stored in database for further investigation. Preprocessing is done by natural language processing pipeline. Literature survey expresses the use of chatbots in medical field as well as various researchers and their experimental analysis of chatbots for various services.

1.2 Literature Survey

This paper explains the systematic literature survey of medical chatbot applications proposed and used by various successful researchers. Taobias Gente et al. proposed a systematic literature review of Medical Chatbot Research from a behavior change perspective in 2020. This paper reviews the published research results relating to details of designing techniques of chatbot. It explains about the design of conversation systems methodology. Emanuela Haller et al. designed a chatbot that can simulates an historical figure. This paper reviews the published research results related developing of architecture chatbots. It explains about the design and implementation NLP and machine learning.

1.3 Methodology

The proposed work consists of the following modules,

- Prepossessing and feature extraction
- Training phase
- Test phase.

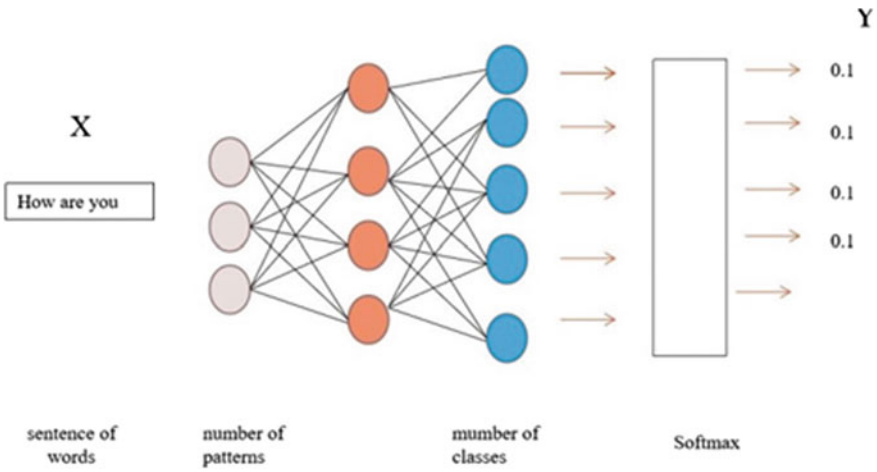


Fig. 1 Feed forward neural network architecture

Prepossessing and feature extraction

It is implemented using the natural language processing pipeline. Tokenization is used for splitting a string into meaningful units, for example, "how are you?" is spitted as ["how," "are," "you," "?"]. Stemming helps to generate the root form of the words. The crude heuristic is used to chops of the ends off words (for example, "organization" is chopped as "organi"). NLP and information retrieval techniques use the concepts of bag of words to convert string to vector. In this model, a text is characterized as the bag of its words by disregarding grammar. This method is commonly used for document classification where the incidence of each word is used as a feature for training a classifier.

Training Model using Deep Learning Algorithms

Figure 1 depicts the feed forward neural network architecture used for training. The step-by-step procedure is given as follows,

- Creation of database which contains information about diseases and its ir symptoms
- User or client sends a message in the form of text or voice.
- NLP uses implicit and explicit expression analysis to identify the similar word patterns in the input data [4]
- NLP converts input to machine understandable format.
- To respond to the user's input, the chatbot creates an answering sentence out of the knowledge stored in the database.

This method does not need manual processing, hence time required for processing the data is very less and provides accurate response [1, 5]. The retrieval-based method

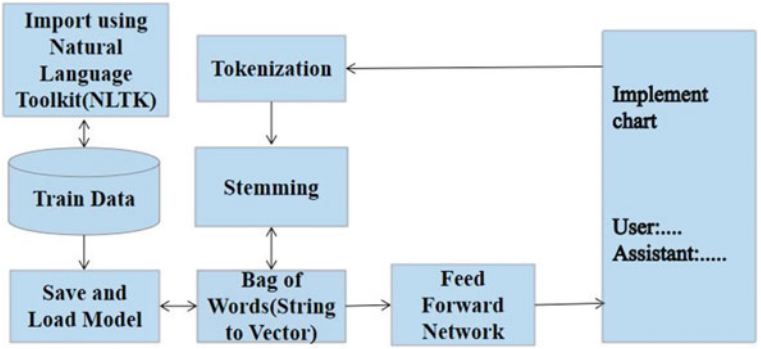


Fig. 2 Proposed architecture diagram

compares queries with message-response pairs and can retrieve sentences with high naturalize and fluency but is usually used in a closed domain.

Test phase

Artificial neural networks (ANNs) and deep learning techniques yield better solutions for many real-world problems such as image recognition, medial data analysis, and speech recognition [6]. NLP plays a vital role in AI field that helps to examine how computer systems can interpret and control [1, 7] text or speech. Most NLP techniques rely on deep learning and machine learning techniques. The proposed system helps to rectify all shortcomings of the existing system and helps users to exploit services maximum and also to provide more advantages than manual methods [8, 9]. In this paper, we propose a general artificial intelligent chatbot agent, and model is built. Deep learning model which analysis the query data from NLP pipeline and provides response to user [10, 11] (Fig. 2).

1.4 Module Description

1.4.1 Preparation of Dataset

Deep learning models are used to train the dataset which is collected from the client. Every new data details collected at the time of application form act as a test dataset.

1.4.2 Preprocess Data

Tokenizing is the first and most basic step in preprocessing. It converts the entire text into small parts like words. Lemmatizing process is applied to convert word into

lemma form. At the end of preprocessing, pickle file is created to store the Python objects.

1.4.3 Creation of Training and Test Data

Training dataset is created from the client input. After preprocessing, it is stored in database for experimental investigation purpose. Output pattern uses bag of words concepts to give response to the client.

1.4.4 Build the Model

- Deep neural network layer which consists of three layers is built using keras sequential API. The model is trained for 200 epochs, and the model is saved as "data.pth"
- Loading the trained model and then use a graphical user interface that will predict the response from the bot.
- The model implements appropriate functions which will identify the class and then retrieve us a random response from the list of responses.

2 Implementation and Testing

2.1 Input and Output

We can test the input by providing details by the user. After that, deep learning model provides the below output Figs. 3, 4, and 5.

2.2 Sample Input and Output

See Fig. 6.

2.3 Advantages of the Chatbot

- Easy to analyze.
- Less time consuming and easy to control.
- External evaluation can be reduced.

Fig. 3 Output image1

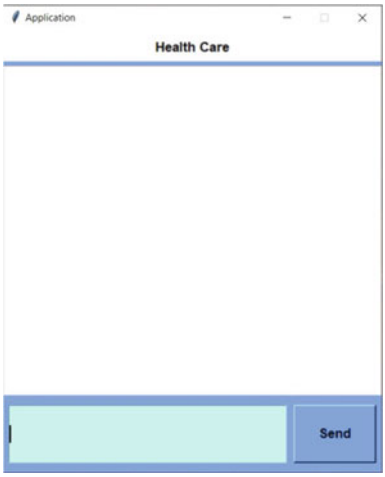
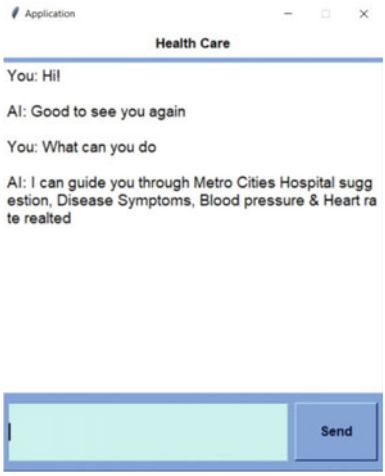


Fig. 4 Output image2



- System can be accessed from anywhere using Internet.
- Fast and more efficient than existing system.

3 Results and Discussion

This work is trained successfully using Python with custom dataset. Data are processed by tokenization, stemming, and bag of words (string to vector) using natural language toolkit (NLTK). Then, training data are created. The PyTorch model and training by feed forward neural network with 3 linear layers and 2 hidden layers

Fig. 5 Output image3

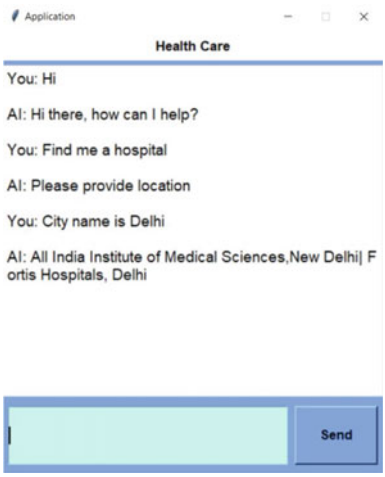


Fig. 6 Test image

```
import random
import json

import torch

from model import NeuralNet
from nltk_utils import bag_of_words, tokenize

device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')

with open('intents.json', 'r') as json_data:
    intents = json.load(json_data)

FILE = "data.pth"
data = torch.load(FILE)

input_size = data["input_size"]
hidden_size = data["hidden_size"]
output_size = data["output_size"]
all_words = data['all_words']
tags = data['tags']
model_state = data["model_state"]

model = NeuralNet(input_size, hidden_size, output_size).to(device)
model.load_state_dict(model_state)
model.eval()

bot_name = "AI"

def get_response(msg):
    sentence = tokenize(msg)
    X = bag_of_words(sentence, all_words)
    X = X.reshape(1, X.shape[0])
    X = torch.from_numpy(X).to(device)

    output = model(X)
    _, predicted = torch.max(output, dim=1)

    tag = tags[predicted.item()]

    probs = torch.softmax(output, dim=1)
    prob = probs[0][predicted.item()]
    if prob.item() > 0.75:
        for intent in intents['intents']:
            if tag == intent["tag"]:
                return random.choice(intent['responses'])

    return "I do not understand..."
```

with Softmax are used for train the data. The system proposed helps to provide health care services for all the people with high efficiency.

4 Conclusion and Future Enhancements

The proposed system helps to provide health care services for all the people with high efficiency. The artificial intelligent agent component of the is receiving increasing attention from researchers. Since these health care services are provide help to user and aid in information. The information provided by chatbot becomes important. This will eradicate the manual method of human interaction to the user in this increasing population which is very tough and hard to provide the service and also the efficiency of services will be very low in manual method. This work is trained successfully using Python with high accuracy. Our work is an initial attempt to develop and provide health care service virtual agent. In future, we can improve it by having dataset and model in the cloud. Hence, the speed of accessing the system could be improved. We use various deep learning models and improve their accuracy and visualize the various parameters that help in improving the accuracy.

References

1. Adamopoulou E, Moussiades L (2020) Chatbots: history, technology, and applications. Mach Learn Appl 2 (2020)
2. Ayanouz S, Abdelhakim BA, Benhmed M (2020) A Smart Chatbot Architecture based NLP and Machine learning for health care assistance. In: The 5th international conference on smart city applications, 2020
3. Siddique S, Chow JC (2021) Machine learning in healthcare communication. Encyclopedia 1:220–239. https://doi.org/10.3390/encyclopedia1010021
4. Gentner T, Neitzel T, Schulze J (2020) A systematic literature review of medical chatbot research from a behaviour change perspective. In: IEEE 44th annual (COMPASAC), 2020
5. Haller E, Rebedea T (2013) Designing a chat-botthat simulates an historical figure. In: IEEE conference publications, July 2013
6. Dzobo K, Adotey S, Thomford NE, Dzobo W (2020) Integrating artificial and human intelligence: a partnership for responsible innovation in biomedical engineering and medicine. Omics J Integr Biol 24:247–263
7. Roos S, Supervisor: Ruth Loc:han (2018) Chatbots in education a passing trend or a valuable pedagogical tool?
8. Yang YC, Al-Garadi MA, Bremer W, Zhu JM, Grande D, Sarker A (2021) Developing an automatic system for classifying chatter about health services on twitter: case study for Medicaid. J Med Internet Res 23(5):e26616
9. Ahmed Z, Mohamed K, Zeeshan S, Dong X (2020) Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database 2020
10. Higashinaka R et al (2014) Towards an open domain conversational system fully based on natural language processing. In: COLING'14, pp 928–939
11. Mustaqeem, Kwon S (2020) A CNN-assisted enhanced audio signal processing for speech emotion recognition. Sensors 20:183

Simulation of WSN-Based Crowd Monitoring and Alert Generation Architecture

Jenny Kasudiya and Ankit Bhavsar

Abstract Simulation tools provide a way to simulate the network under different conditions, test the network, and analyze the results produced. One such simulation tool is network simulator (NS2). This paper presents the preliminary performance analysis of the WSN-based crowd monitoring architecture with the help of NS2. The architecture tries to reduce the occurrences of casualty by monitoring the individuals at crowd gatherings, keeping a track of the crowd in a sector, and disseminating the crowd to neighbouring sectors when the need arises. NS2 tries to simulate the architecture, and the generated results are analyzed based on various criteria like throughput, end-to-end delay, energy consumption, and many others.

Keywords Wireless sensor network (WSN) · Simulation · Network simulator2 (NS2) · Crowd · Crowd monitoring

1 Introduction

1.1 WSN-Based Crowd Monitoring Architecture

Wireless sensor network (WSN) is a network of geographically distributed sensors. The sensors are used to measure various atmospheric parameters [1]. The readings of these sensors are accumulated at a central location and necessary actions taken accordingly. A sensor is made up of sensing unit, processing unit, power unit, and communication line [2, 3].

Crowd is large mass of individuals coming together for an event in an absolutely unordered manner. When the number of individuals gathering for the event crosses the expected count, it results into chaos, pushing, trampling, and thereby ruining the

J. Kasudiya (✉) · A. Bhavsar
GLS University, Ahmedabad, Gujarat, India
e-mail: jenny.kasudiya@glsuniversity.ac.in

A. Bhavsar
e-mail: ankit.bhavsar@glsuniversity.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_3

property. At times, it also risks the life of individuals coming together [4]. To reduce the chaos and loss of lives at such gatherings, a WSN-based architecture has been proposed in [4] for crowd monitoring and alert generation.

The architecture shows various sensors that are deployed to monitor the crowd, their movements, and their behaviour. During abnormality, the sensor-based system will get activated and guide the crowd for safe evacuation. The architecture tries to extend the traditional setup for crowd gathering and management to reduce human accidental death as much as possible [4].

The architecture proposed in [4] ensures the safety and security of the pilgrims during normal as well as emergency situations of crowd gathering. The sector contains entities like centralized sensing and monitoring unit (CSMU), security guard, paramedic, smart navigation screens (LCD), pilgrim wrist band, and emergency exit pathways. CSMU acts as a crowd monitoring and controlling unit within a given sector and is responsible for various activities like keeping track of crowd count in a sector, communicating with the control room, opening/closing of barricades, and many more. In case of the crowd in a sector is more than expected, the sensor-based system with the help of pathways disseminate individuals to the nearby sector by opening of the barricades. The use of WSN-based architecture, we try to reduce the number of casualties occurring at mass gatherings [4].

1.2 Network Simulator (NS2)

Simulation is the process of designing and developing hypothetical model of a real-time system that can be useful for understanding the behaviour of the network, its interconnection, and the communication between various entities under consideration. Developing and testing a real-time network are pretty expensive. Hence, simulation tools provide an easy solution to testing your network [5–7].

A wide variety of open-source network simulators are available today, one of them popularly known as network simulator-2(NS2). NS2 is an object-oriented discrete event simulator. It uses programming languages like C++ and tool command language with object oriented extension (OTCL). We can generate different topologies, simulate routing protocols like TCP and UDP, can select different networking devices like bridge, hub, router, and switch, and also have wired/wireless communication media. One of the advantages of using NS2 is that complex network scenarios can be tested, and the results can be easily generated [5–9].

NS2 also provides network animator (NAM), allowing the user to capture animation traces of network simulation and packet exchange. NAM helps in graphical representation of the network elements and can be used to visualize the progression of packets when the simulation starts. There are various options to start, pause, and stop the animation in NAM window. NS2 also makes it possible to calculate the performance of the network using AWK and PERL scripting languages. It allows measuring various parameters like end-to-end delay, throughput, and packet drop ratio using the scripting languages [10, 11].

This paper is divided into four sections. Section 2 provides details of the simulation scenario. Section 3 shows the outcome of the simulation along with the performance comparison followed by conclusion in Sect. 4.

2 Simulation Scenario

2.1 Simulation Process

Simulation is the process of developing an exact prototype/replica of the object under study. The study of the object involves basic understanding of how the system works. Such a simulation can also be useful to study a network. Practically, buying all the components and implementing a network are very expensive. Hence, network simulation is done to test the computerized network under various conditions which are hard to simulate in a live network [5, 7, 8].

The architecture proposed in [4] is composed of many sectors where each sector contains a CSMU unit which is responsible for managing the crowd within the sector by keeping a track of the crowd count. As soon as the number of people gathering in a sector crosses threshold, CSMU informs the control room about it. On receiving acknowledgement from the control room, CSMU disperses the crowd to the nearby sectors. In this manner, the crowd in a sector is managed to avoid chaos and occurrence of accidents. In this paper, we try to simulate the architecture proposed through the use of NS2 simulation tool. Let us go through the steps followed in the simulation model for the data packet transmission using WSN:

1. Two sectors are created identified as sector 1 and sector 2.
2. Several nodes are created in sector 1.
3. Each sector will have a CSMU unit denoted in red colour.
4. As the number of nodes in sector 1 increases beyond the threshold, process of crowd dispersion needs to start.
5. A request message for crowd dispersion is sent to the control room.

On receiving an acknowledgement from the control room, few nodes are dispersed from sector 1 towards sector 2.

2.2 Simulation Parameters

For implementing the proposed architecture in [4], a simulation of this entire scenario is done in NS2. We have created 22 sensor nodes and two CSMU nodes (represented by sector 1 and sector 2 here) in NS2. The sensor nodes correspond to the pilgrims moving inside the sector, and the server nodes represent the CSMU unit. We would be using wireless channel for simulation. It uses ad hoc distance vector (AODV)

routing protocol and unreliable user datagram protocol (UDP). The simulation time is set to 25 s. Several other parameters set up initially for the simulation are shown in Table 1.

The above figures depict the simulation scenario. Figure 1 shows the initial state of the network created in NS2 where the green colour denotes the sensor nodes, and red colour depicts the CSMU unit. The same figure also shows that sector 1 is getting

Table 1 Initial simulation parameters

S. No	Simulation parameters	Values
1	NS version	NS 2.35
2	Channel	Wireless channel
3	No. of nodes	22
4	Type of MAC	802.11
5	Antenna model	Omni directional
6	Type of routing protocol	AODV
7	Internet protocol type	UDP
8	Area of simulation	1800*840
9	Propagation	Two ray ground
10	Interface queue	Drop tail/priority queue
11	Traffic type	cbr
12	Max packet size	1000
13	Simulation time	25 s

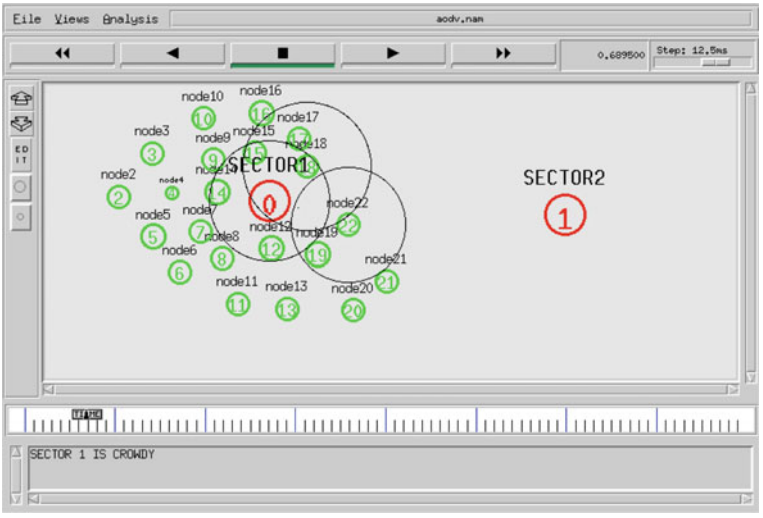


Fig. 1 Simulation scenario: sector 1 gets overcrowded

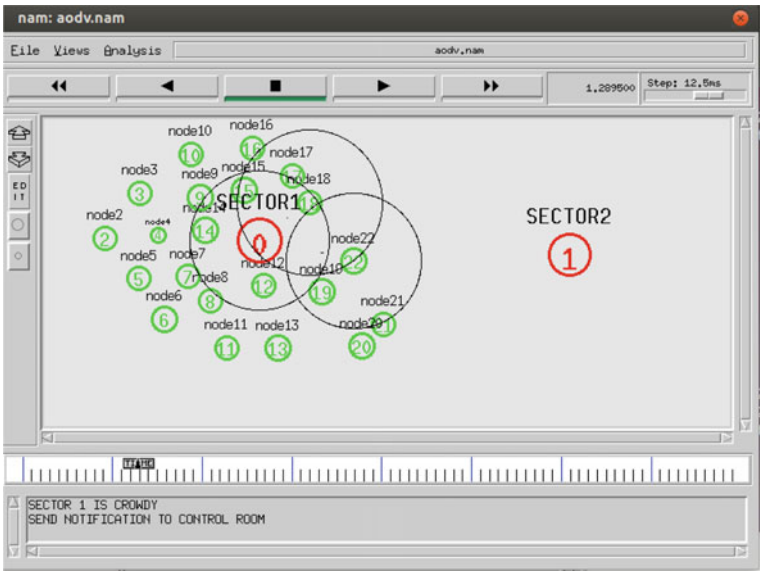


Fig. 2 Simulation scenario: sector 1 CSMU sends a notification to the control room

over crowded due to the increase in the number of nodes moving inside it. As sector 1 is getting overcrowded, the CSMU unit in sector 1 sends a notification to the control room. Figure 2 shows a notification is being sent to the control room.

Figure 3 shows that sector 1 on receiving an acknowledgement from CSMU, nodes 17, 18, 19, 20, 21, and 22 start moving towards sector 2. The entire simulation scenario shows how the crowd in a sector is dispersed to the nearby sector in order to minimize any casualty.

3 Simulation Outcome

3.1 Result Analysis Based on Various Parameters

The NS2 simulation files are written in tool command language (TCL) with.tcl extension. On executing.tcl files, a trace file gets generated. The trace file contains the overall network information. Apart from simulation, NS2 provides various ways to an- alyze the performance of the simulated network. To measure various parameters, AWK programming language is used. AWK scripts can efficiently process the data from the trace files [8, 11]. Using the AWK scripts, we have measured several parameters for the simulated network. Let us see the details of each one:

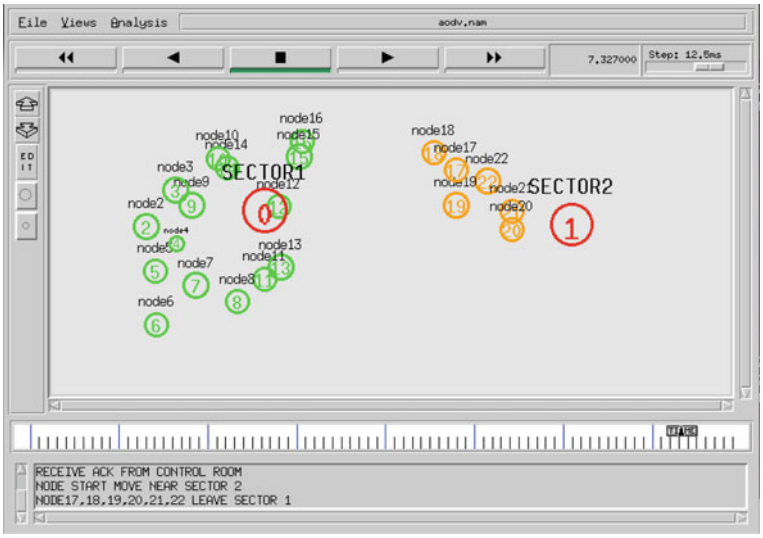


Fig. 3 Simulation scenario: nodes are moving towards sector 2

3.1.1 Probability Density Function (PDF)

The ratio of number of packets received to the number of packets sent is known as probability density function [11, 12]. It is measured in bits per second (bps). Figure 4 shows PDF for the 22-node network.

```
jenny@ubuntu: ~/Documents/Code_Files
File  Edit  View  Search  Terminal  Help
 0.012500000000000001 1
Please use this format in the future.
v -t <time> -e <tcl expression>

^C
jenny@ubuntu:~/Documents/Code_Files$ awk -f pdf.awk aodv.tr
cbr s:606 r:507, r/s Ratio:0.8366, f:0
jenny@ubuntu:~/Documents/Code_Files$ awk -f energy.awk aodv.tr
node 0
node 1
node 2
node 3
node 4
node 5
node 6
node 7
node 8
node 9
====================
average 72
====================
total energy 720
====================
jenny@ubuntu:~/Documents/Code_Files$
```

Fig. 4 Probability density function and energy consumption of 22-node network

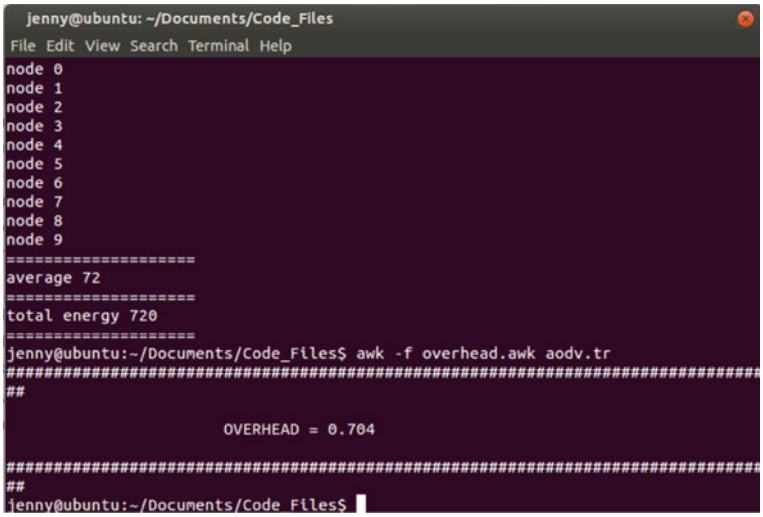


Fig. 5 Overhead of 22-node network

3.1.2 Energy Consumption

The life span of a node in WSN depends on the energy consumption of the sensor nodes which when reaches zero, the node may no longer be able to send/receive packets [11, 13]. Hence as the energy consumption of the nodes is minimized, the network life time increases. The unit of energy consumption is Joule. Figure 4 shows energy consumption for the 22-node network.

3.1.3 Overhead

Routing packets are sent across to identify the communication link failures, device failure, and many more such issues in the network [14]. Lesser the number of packets sent for route discovery indicates better the state of the network. Network overhead is measured in bps. Figure 5 shows overhead for the 22-node network.

3.1.4 Throughput

Number of packets that can be successfully delivered in unit time can be called as throughput of the network [12–14]. Higher the value of throughput, better the performance of the network. It is measured in bps. Figure 6 shows throughput for the 22-node network.

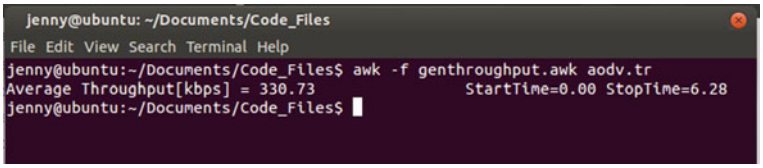


Fig. 6 Throughput of 22-node network

Table 2 Performance comparison of 500 and 1000 nodes based on various criteria

	No. of nodes	
Criteria	500	1000
Probability density function (bps)	0.2547	0.4796
Energy consumption (Joule)	1.141	0.698
Throughput (bps)	129.89	220.61
Overhead (bps)	116.766	80.536
e2e delay (ms)	194.02	82.88

3.1.5 End-to-End Delay

The average time it takes to route a data packet from the source to the destination is known as end-to-end delay [12–14]. Lower the value of delay, better is the performance of the system. Delay is generally measured in milli-seconds.

3.2 Performance Comparison

With the same initial parameters set, Table 2 shows that by increasing the number of nodes to 500 and 1000, we can see the performance of our simulated network. The values in Table 2 clearly depict that there is a decrease in the energy consumption, overhead, and delay resulting into a considerable increase in the throughput of the network. Henceforth from chart in Fig. 7 shows that increase in the number of nodes will give a better performance of the network.

4 Conclusion

This paper shows the implementation of WSN-based architecture for crowd monitoring and alert generation with the help of NS2 simulator. Upon simulation of the architecture with 500 and 1000 nodes, the performance of the network significantly improves when throughput parameter is considered. It is also observed that with the

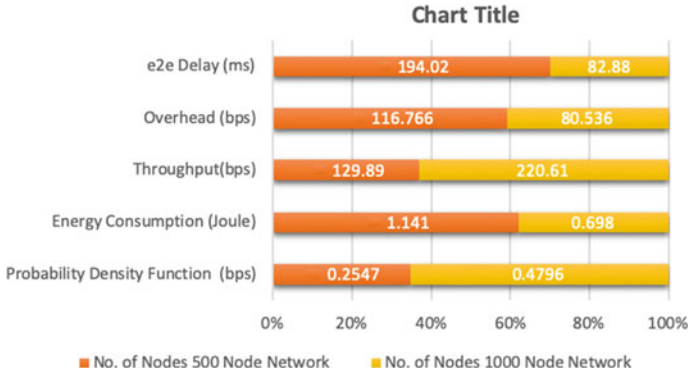


Fig. 7 Performance comparison chart of 500 and 1000 nodes based on various criteria

increase in the number of nodes, there is a reduce in the delay. The simulation gives encouraging results for the WSN-based architecture for crowd monitoring and alert generation.

References

1. Kasudiya J, Bhavsar A (2021) Communication protocols for WSN based crowd monitoring architecture. High Technol Lett 27(1). ISSN: 1006-6748
2. Rao KR, Tirumala Rao SN, Reddy PC (2016) A review on wireless sensors network. Int J Comput Trends Technol 40(2). ISSN: 2231-2803
3. Kazem S, Minoli D, Znati T (2007) Wireless sensor networks: technology, protocols, and applications. Wiley
4. Kasudiya J, Bhavsar A, Arolkar H (2019) Wireless sensor network: a possible solution for crowd management. In: Springer organized 2nd international conference on smart IoT systems: innovations in computing, Jaipur
5. Siraj S, Kumar Gupta A, Badgujar R (2012) Network simulation tools survey. Int J Adv Res Comput Commun Eng 1(4)
6. Chaudhary R, Sethi S, Keshari R, Goel S (2012) A study of comparison of network simulator -3 and network simulator -2. Int J Comput Sci Inf Technol 3(1):3085–3092 (2012)
7. Pandey S, Tyagi V (2013) Performance analysis of wired and wireless network using NS2 simulator. Int J Comput Appl 72(21)
8. Khandelwal A (2018) A study on network simulator 2. Int J Curr Res Life Sci 07(02):1036–1039. ISSN: 2319-9490
9. Pan J, Jain R (2008) A survey of network simulation tools: current status and future developments. Email: jp10@ cse. wustl. edu 2(4):45
10. Arvind T (2016) A comparative study of various network simulation tools. Int J Comput Sci Eng Technol 7(08):374–378. ISSN: 2229-3345
11. Gautam G, Sen SB Design and Simulation of Wireless Sensor Network in NS2. In: International Journal of Computer Applications (0975 – 8887), Volume 113 – No. 16. March (2015).
12. Manikandan S, Manikanda Kumaran K (2014) Performance analysis of mobile Ad-Hoc network routing protocols using network simulator. COMPUSOFT I J Adv Comput Technol 3(6). ISSN:2320-0790

13. Bassi J, Aulakh T. A novel approach for reducing energy consumption and increasing throughput in wireless sensor network using network simulator 2. Int J Comput Appl 0975–8887
14. Rampurkar K, Lavande N, Shilgire S, Mane SN (2017) Study of routing overhead and its protocols. Int J Adv Eng Manage 2(2):52–55

Real-Time Assessment of Live Feeds in Big Data

Amol Bhagat and Makrand Deshpande

Abstract An extremely challenging proposition is constructing a functional, well-designed, and reliable big data application that supplies to a variety of end-user latency necessities. Let alone constructing applications that work for the problem at hand, it can be discouraging plentiful to just keep up with the fast pace of technology innovation happening in this space. But, there are assured high-level architectural concepts that can help to imagine how diverse types of applications fit into the big data architecture and how several of these technologies are renovating the existing enterprise software scene. Big data analytics is used to gather valuable findings, identify trends and detect patterns from the ocean of information. Proficient operators say its vital to assess the prospective business significance that big data software can offer. This paper presents a study of the lambda architecture enterprise design to build a data-handling back-end on cloud, assuring intense, high throughput, and dense data demand provided as services. This paper also covers the advantages proposed by the lambda architecture and how effectively it can be used across wide range of use cases across the industry. Based on the experience gathered during this research—this paper also discusses the challenges those would be posed by lambda architecture.

Keywords Big data · Hadoop · Lambda architecture · Live feeds · Real-time assessment · Splout · Storm · Trident · Tweets

1 Introduction

Big data is a term for data that are so multifaceted or huge that traditional data processing system is insufficient to deal with. Challenges are storage, capture, analysis, search, curation, sharing, visualization, transfer, querying, information privacy,

A. Bhagat (✉)
Department of Computer Science and Engineering, Prof Ram Meghe College of Engineering and Management, Badnera, Amravati, Maharashtra 444701, India
e-mail: amol.bhagat84@gmail.com

M. Deshpande
Citicorp Services India Limited, Pune, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_4

and updating. The word "big data" frequently denotes the usage of user behaviour analytics, predictive analytics, or additional advanced data analytics techniques that mine value from data. Data sets develop swiftly—in component because they are more and more amassed by means of reasonably priced and several data-sensing mobile devices, aerial (faraway sensing), software programme logs, microphones, cameras, wireless sensor networks and radio-frequency identity (RFID) readers. The sectors technological in step with capita potential to shop information has kind of doubled every forty months since the Nineteen Eighties; as of 2012, each day 2.5 Exabyte (2.5×1018) of information are generated.

Real-time analytics is the usage of or the capacity to use, records and related resources as quickly because the information enters the device. Real-time analytics is likewise known as dynamic analysis, real-time analysis, real-time records integration, and real-time intelligence. Real-time big facts is not just a manner for storing petabytes or Exabyte of statistics in a statistics warehouse, it is about the capacity to make better decisions and take significant actions at the proper time. It is about detecting fraud whilst someone is swiping a credit score card, or triggering an offer whilst a client is standing on a checkout line, or setting an advert on Internet site whilst a person is studying a selected article. It is about combining and analysing records so you can take the appropriate action, at the proper time, and on the proper vicinity. Figure 1 shows different opinions of the researchers in the Big Data Analytics domain.

Large statistics analytics cannot simplest assist corporations create new efficiencies, make better choices and facet into real-time processing, it could additionally inspire enterprise transformation. One characteristic of the massive information revolution is the acknowledgement that single facts-base management gadget structure cannot meet all desires. As an instance, when a device is needed that could reliably

Fig. 1 What research companies has to say on real-time analytics

procedure ACID (Atomic, regular, independent, long lasting) transactions then relational database system is the way to go. If, however, whilst want is a device that may keep availability on heavy information load whilst there may be two cut up network outage, then it requires a non-relational regular system. Of route, for low-budget storage and processing of big amounts of raw unstructured data, Hadoop is usually favoured.

Huge records architectures regularly boil all the way down to actual time or batch with typhoon or Hadoop map-lessen. The amalgam lambda structure demanding situations this separation. The lambda structure attempts to offer a mixture of technology that together can provide the characteristics of an Internet-scale a web scale device that may fulfil requirements for availability, maintainability, and fault tolerance. It also enables a hybrid structure—which merges batch layer—which computes perspectives on collected statistics and actual time layer—which is parallel speed processing layer to shut the distance by means of continuously processing the maximum current facts.

Lambda structure is a useful framework to reflect on consideration on designing large information applications. Nathan Marz designed this familiar structure addressing commonplace requirements for huge records primarily based on his experience working on distributed information seasoned-cessing structures at Twitter. A number of the important thing requirements in constructing this architecture comprise: Fault tolerance in opposition to hardware disasters and human inaccuracies support for a spread of usage instances that consist of short latency querying in addition to updates Linear scale-out competencies that means that throwing extra machines at the trouble must help with getting the job finished extensibility so that the machine is manageable and can accommodate more recent capabilities easily.

The impression of the lambda structure is straightforward and two-fold. Initially, a batch layer calculates views on your accumulated data and recurrences the technique whilst it is executed to endlessness. Its yield is continually previous by the point its far obtainable seeing that new statistics has been acquired in the period in-between. Subsequent, a parallel pace handling layer ends this gap by way of continually handling the latest data in immediate real-time. Because of their popularity, Hadoop and typhoon are sturdy candidates for the velocity and batch layer of recent information designs. Though, any technology, which include a company New York's legacy software demonstrating a speed or batch layer function, could fulfil any function. Figure 2 shows the detailed lambda architecture.

2 Literature Review

Manny Punnets of online publicity platform suppliers Altitude Digital and Share through are together beating Apache Spark's stream handling competencies to provision extra real-time investigation of ad data. Many data coursing applications do not comprise enormous amounts of information. Ed Burns from TechTarget has cautioned tech savvy people—a numeral of fables approximately massive records

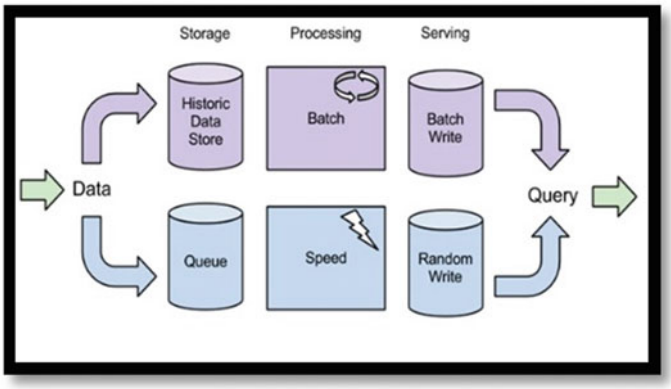


Fig. 2 Detailed lambda architecture

have flourished in latest years. Do not let those not unusual misunderstandings destroy your analytics undertaking. Ed Burns has also highlighted the importance to gaze at the big picture before agreeing any big data tool [1].

Usama Fayyad, a highlighted narrator at Netezza's Enzee Universe consumer convention, got here armed with facts on how the forthcoming of "large statistics" lies in analytics. The three V's, or what IBM and the analyst firm Gartner Inc. use to define "big data," are testing the limits of traditional data warehouses and necessary functions such as extract, transform and load (ETL) capabilities, conferring to Usama Fayyad, CEO of the data consulting, strategy, and technology firm Open Insights and previous executive vice president and chief data officer of research and strategic data solutions at Yahoo [2].

In an investigation piloted by Capgemini Consulting ultimate November, most effective 35% of the 226 analytics, IT and commercial enterprise specialists who replied defined their huge data projects as a hit or precise effective. One of the massive reasons, in line with a record on the investigation, is that maximum agencies "are a ways from being capable to use [big] records correctly". For example, best 35% of the respondents believed their companies had sturdy procedures for shooting, curating, authenticating and preserving information, even as 79% said that they had but to fully combine all of their statistics assets [3–5].

The fortunate organization Inc. attempts to acquire price from its analytics determinations via retaining possessions in framework. Lucky group gathers inner data on month-to-month income, product income and what pages guests are viewing at on its sites, which comprise StyleMint.com and JewelMint.com. The corporation also collects purchaser facts, together with what products humans purchase and how far they spend. It practises Pentaho's data analytics and integration tools to capture the information into a MySQL database [6–8] and then analyse it.

Machine learning algorithms are also attempted in real-time stream analysis. Once such concept is presented by Ms D Jayathi in their work of—A Framework for Real-time Streaming Analytics using Machine Learning Approach. The cost effectiveness

of the sensor data analysis solution based on big data lambda architecture is analysed in [9–13].

The increasing range of ordinary embedded devices which are intersected over the net ends in the requisite of new software programme answers for coping with them in an effectual, ascendable, and clever way. In addition, such devices yield a large quantity of information, inflicting the large facts trouble that want to be saved and processed. This emerging state of affairs, called Internet of things (IoT), raises fundamental demanding situations: massive-scale smart environs administration and big statistics garage/analytics. Both challenges are addressed in [14, 15], proposing AllJoyn lambda, a software programme answer incorporating AllJoyn in the lambda architecture utilized for big facts storage and analytics. It addressed how may be conceivable to accomplish embedded gadgets in a smart style seasoned-cessing together batch and real-time records. A technique of enhancing speed of judgement making by means of analysing real-time streams for operational business intelligence with traditional gadget and to present speedy outcomes for improvised decision making is presented in [16].

A new algorithm-based totally on a weighted projection quantiles for immediate and frugal real-time quantile estimation of big sized excessive dimensional facts clouds is presented in [17]. A projection quantile regression set of rules for excessive dimensional records is presented. A fast algorithm for computing the depth of a factor or a brand new observation in terms of any high-dimensional information cloud is presented and a ranking machine for multivariate data is proposed. A real-time rapid tracking scheme is briefly described just like statistical method tracking, for actionable analytics with huge statistics.

Today's network forensics is belligerent to convey collectively widespread capacities of dissimilar information from all styles of logs and network site visitors, and to examine the large volume of statistics in actual time. The layout, implementation, and evaluation of an actual time dataflow manner platform the use of an amalgamation of open supply gear that captures, preserves and examine excessive volume of network sports and logs to in addition support community forensic is presented in [18]. Proposed platform in [9, 18, 19] is now being utilized in a huge company in China and has been ascertained to accomplish high performance and can be utilized in real-time with the maximum throughput of around 5 Gb/s. Figure 3 shows the real-time analysis proposed in [20, 21].

Twitter has obtained an awful lot interest lately. A crucial function of Twitter is its real-time nature. The real-time communication of activities together with earthquakes is investigated in [21] the use of Twitter and proposed a set of rules to monitor tweets and to come across a goal occasion. To come across a target occasion, devised a classifier of tweets based on functions which includes the key phrases in a tweet, the range of words, and their situation. Each Twitter user is regarded as a sensor and applies particle filtering, which might be broadly utilized for area assessment. The particle filter out works higher than other comparable techniques for assessing the locations of goal occasions. As a utility, an earthquake reporting gadget is advanced for use in Japan. Due to the numerous earthquakes and the large variety of Twitter users, we will locate an earthquake with excessive probability (93% of earthquakes

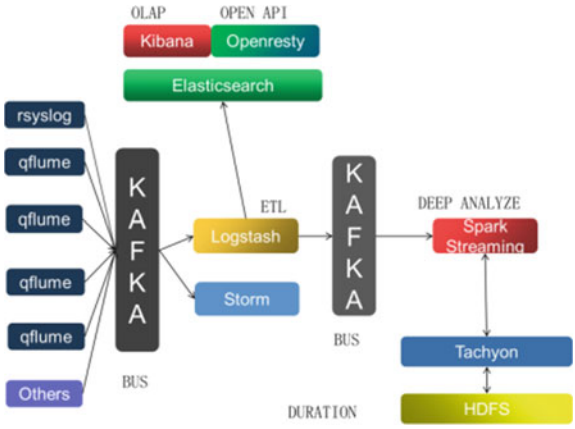


Fig. 3 Real-time analysis

of Japan Meteorological organization (JMA) seismic depth scale three or greater are spotted) merely with the aid of tracking tweets. Machine outlined in [21], claimed to detect earthquakes directly and notification is introduced much quicker than JMA broadcast bulletins.

3 Proposed Methodology for Real-Time Assessment

The lambda architecture centrally gets records and does as low as probable seasoned-cessing before copying and splitting the facts stream to the actual time and batch layer. The batch layer gathers the information in an information sink like HDFS or S3 in its raw shape. Hadoop jobs frequently technique the statistics and write the end outcome to a statistics store. The main components of lambda architecture are specified below:

New data: All information coming into the machine is transmitted to each the batch layer and the speed layer for handling. **Batch layer**: this sediment has capabilities: (i) handling the master dataset, an incontrovertible, append-handiest set of uncooked statistics, and (ii) to pre-compute arbitrary query features, referred to as batch perspectives. Hadoop's HDFS is generally used to shop the master dataset and perform the computation of the batch interpretations using MapReduce. **Serving layer**: this residue indexes the batch views so they can be interrogated in ad hoc with low-latency. To put into effect the serving layer, usually technologies inclusive of Apache HBase or ElephantDB are applied. The Apache Drill venture offers the capability to execute full ANSI sq. 2003 queries in opposition to batch views.

Velocity layer/Real-time layer: this residue compensates for the excessive latency of apprises to the serving layer, because of the batch layer. The usage of rapid and incremental algorithms, the speed layer offers with latest facts only. Storm is often used to put in force this accretion.

Queries: closing however no longer least, any incoming question may be responded through merging effects from batch views and real-time views.

The batch perspectives have to be recomputed from scratch from the entire grasp dataset. You may think that that is a horrific concept and virtually its extra performant to put into effect incremental MapReduce algorithms to boom the frequency of the batch perspectives. even though by way of doing this it is far viable to alternate performance for human fault tolerance, because whilst the use of an incremental algorithm it is a whole lot more hard to repair any problems in the perspectives. To optimize the facts set and enhance question efficiency, diverse rollup and aggregation techniques are executed on uncooked statistics [6], even as estimation techniques are employed to in addition reduce computation fees [7]. Whilst expensive full recomputation is needed for fault tolerance, incremental computation algorithms may be selectively added to boom efficiency, and techniques together with partial computation and useful resource-usage optimizations can effectively assist lower latency [9].

4 Proposed System Implementation Details and Evaluations

In this paper calculating the variety of occurrences of hashtags in tweets, assembled by way of date, and serve the statistics as a faraway provider is applied, as an example on the way to populate timelines in an Internet site/cell app (e.g. list out the progression of indications for hashtag "California" for the beyond 10 days). The necessities for the answer are: it need to scale (to method billions of tweets.), and it need to be capable to assist low-latency requests to doubtlessly loads of simultaneous customers enquiring for timelines. By means of Hadoop to keep the tweets and an easy Hive question for assemblage via hashtag and date appears worthy sufficient for computing the counts. But, we also want to feature real-time to the device: we need to have the real quantity of appearances for hashtags reorganized for today in second's time. We need to place the Hadoop counts in a few honestly rapid datastore for being capable of question them.

The solution proposed is to apply "lambda architecture" and put into effect an actual time layer utilizing Trident, that's an API on pinnacle of typhoon that simplifies constructing real-time topologies and saving chronic kingdom derived from them. For helping the batch layer Splout SQL is used that is an excessive-performant SQL examine-simplest facts accumulation that may serve and pull datasets from Hadoop very competently. Splout is rapid similar ElephantDB but it also permits us to accomplish SQL queries. The use of SQL for aiding the batch layer is suitable as we might need to break-down the counts by week, day, hour, or any random date period. Trident is used to implement the remote service using its DRPC abilities. Trident itself will question both the real-time layer and the batch layer and combine the results. Ubuntu Linux 16.0, Cloudera Cloud VM 4.5, Hadoop 2.7.3, Apache

Spark 2.1.0, and Apache Kafka 0.10.1.1 [22–24] are used for the development and deployment of proposed system for real-time data analysis.

Splout SQL is a database used for check cause that could pull data from Hadoop very efficaciously. It is a divided, read-most effective, tremendously performant sql database for Hadoop-scale datasets. All of the portions glued laidback to procedure a Lambda. Tweets are served into the gadget, via a kafka topic. The topic is configured to just accept the incoming feeds from the twitter. A Trident flow keeps them into Hadoop (HDFS) and handles them in real-time for growing an in-memory state with the calculations through date. Afterwards, Splout SQL's [10] is used as command-line or API gear for indexing the record and installing it to a Splout SQL cluster, with a purpose to be capable to serve all the information appealing speedy. Then, a 2nd Trident circulate (DRPC) can be utilized to serve timeline queries, and this move will query both the batch layer (thru Splout SQL) and the real-time layer (through the first circulation's memory kingdom), and blend the outcomes into a sole timeline reaction. That is how, conceptually, the overall architecture seems like as shown in Fig. 4.

For a dataset that can be GBytes or even TBytes worth of data to serve low-latency queries under high throughput (multiple concurrent users) is one of the challenges of the architecture. Since nearly entire statistics are calculated in Hadoop, some

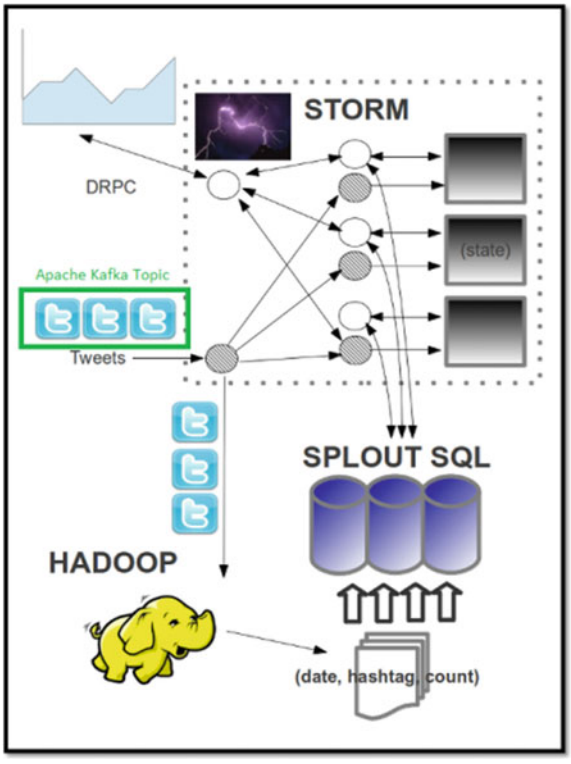


Fig. 4 Proposed system implementation architecture

kind of association between it and a fast database is required. It is not possible directly query files produced from Hadoop. Equipment like Impala or Drill permit you to interactively discover the dataset, but you cannot just place them in facade of a web server. Much needed B-Trees or indexes that can create lookups truly fast. It is possible to evenhanded dump the statistics to a key/value accumulation similar ElephantDB or Voldemort, keyed by hashtag. But to provide daily, weekly or monthly timelines further more to cluster by hour to display the greatest lively hours—Consuming a fast key/value accumulation for Hadoop continually means you want to pre-calculate entirety you are going to necessity for your interpretation, which is not permanently that sustaining. Splout SQL is typically more suitable as it permits for flexible, richer querying of pre-aggregated data. Splout SQL's command-line tools for deploying and indexing gaze like as shown below.

```
hadoop jar splout-hadoop-*-hadoop.jar simple-generate -i sample-hashtags -o out-hashtags -pby hashtag -p
2 -s "label:string,date:string,count:int,hashtag:string" --index "hashtag,date" -t hashtags -tb hashtags
hadoop jar splout-hadoop-*-hadoop.jar deploy -q http://localhost:6615 -root out-hashtags -ts hashtags
```

In batch layer all tweets are appended in the HDFS and intermittently run certain modest procedure that aggregates them by hour date. The dataset small_samples_hashtags is used for processing. The yield of the batch layer should look similar a tabularized text file with hashtags counts as shown below.

```
hashtag    2021082828 2   calirfonia
hashtag    2021102824 2   caliroadtrip
hashtag    2021102926 1   caliroadtrip
hashtag    2021081819 2   caliroots
hashtag    2021093810 2   caliroots
```

Output will look like.

```
...
Result for hashtag 'california' -> [[{"20091022":115,"20091023":115,"20091024":158,"20091025":19}]]
Result for hashtag 'california' -> [[{"20091022":115,"20091023":115,"20091024":158,"20091025":19,"20130123":76}]]
Result for hashtag 'california' -> [[{"20091022":115,"20091023":115,"20091024":158,"20091025":19,"20130123":136}]]
Result for hashtag 'california' -> [[{"20091022":115,"20091023":115,"20091024":158,"20091025":19,"20130123":192}]]
Result for hashtag 'california' -> [[{"20091022":115,"20091023":115,"20091024":158,"20091025":19,"20130123":232}]]
Result for hashtag 'california' -> [[{"20091022":115,"20091023":115,"20091024":158,"20091025":19,"20130123":286}]]
...
```

5 Conclusion

In the lambda structure, real-time is the ability to method the delta of information that has been captured after the start of the batch layers contemporary generation and now. The addition of some other layer to structure has foremost benefits. Firstly, statistics can (traditionally) be treated with high precision and worried algorithms without trailing quick-time period data, indicators, and insights provided with the aid

of the real-time layer. Then, the addition of a layer is offset via vividly decreasing the arbitrary write storage necessities. The batch write garage gives also the selection to switch information at pre-defined times and version information.

Ultimately and prominently, the addition of the records sink of uncooked data gives the choice to pull through from human errors, i.e. installing insects which write inaccurate accumulated information from which other architectures cannot get better. Every other option is to looking back beautify information withdrawal or learning algorithms and relate them at the complete of the ancient dataset. This is extraordinarily beneficial in start-up and agile environments in which MVPs push what can be performed down the song. Lambda structure duplicates incoming facts. Complete recomputation vs. partial recomputation: in a few instances, its miles really worth the use of Bloom filters to circumvent whole recomputation. Recomputational algorithms versus incremental algorithms: there may be a massive temptation to use incre-intellectual algorithms, however according to a guiding principle we ought to use recomputational algorithms even though it makes it more difficult to achieve the identical end result. Additive algorithms vs. approximation algorithms: Lambda structure works properly with additive algorithms. Consequently that is another case we want to do not forget the use of approximation algorithms, as an example, HyperLogLog for a count-distinct problem, and many others.

Criticism of lambda architecture has targeted on its inherent complexity and its limiting influence. The batch and streaming aspects each require a distinctive code base that have to be maintained and stored in sync so that processed records produces the same result from each paths. yet attempting to summary the code bases right into a single framework puts the various specialized tools inside the batch and actual time ecosystems out of attain [4].

In a technical dialogue over the merits of employing a pure streaming approach, it turned into mentioned that using a bendy streaming framework inclusive of Apache Samza may want to offer some of the equal advantages as batch processing without the latency [5]. The sort of streaming framework may want to permit for gathering and processing arbitrarily big home windows of information, accommodate blockading, and deal with nation. With things like Yahoo's hurricane-Hadoop for YARN, and big facts gear being these days simpler and simple to deploy and manage (e.g. Cloudera manager or Apache Ambari creating existence lots less complicated) it turns into much less of an ache to set up an open-source massive records structure, put in force batch processing and collocate it with real-time handling. What's extra, as said, many cases will still be without problems solvable just by batch processing. As we have seen, more modern real-time structures may already offer re-processing or re-indexing semantics under. Due to the various likely use cases, serving requirements, and facts natures, its miles doubtful whether frameworks similar Sum-mingbird or Lambdoop can be key to enforcing lambda architectures.

References

1. Lynch C (2014) Community cleverness required Nature 455
2. Reichman OJ, Jones MB, Schildhauer MP (2011) Challenges and opportunities of open data in ecology. Science 331(6018):703–705
3. Brueckner R (2008) Parallel programming in the age of big data, Gigaom Blog
4. Jay K (2014) Questioning the lambda architecture. In: Oreilly internet archive book
5. Mishra P (Nov 2017) India's big data hunt for cures to mental, ageing-related diseases. *Hackers News*
6. Ahmed E (2017) Mubashir Husain Rehmani, Mobile edge computing: opportunities, solutions, and challenges. Futur Gener Comput Syst 70:59–63
7. Nguyen DT, Jung JE (2017) Real-time event detection for online behavioral analysis of big social data. Futur Gener Comput Syst 66:137–145
8. Apiletti D, Baralis E, Cerquitelli T, Garza P, Pulvirenti F, Venturini L (2017) Frequent itemsets mining for big data: a comparative analysis. Big Data Res 9:67–83
9. Subbu KP, Vasilakos AV (2017) Big data for context aware computing—perspectives and challenges. Big Data Res 10:33–43
10. Splout SQL's architecture. http://sploutsql.com/architecture.html
11. Central Repository Maven. http://search.maven.org/#browse%7C-1223220252
12. Tennant M, Stahl F, Rana O, Gomes JB (2017) Scalable real-time classification of data streams with concept drift. Future Gener Comput Syst 75:187–199
13. Georg Krempl, Indre Žliobaite, "Open challenges for data stream mining research", ACM SIGKDD Explorations—Special issue on big data archive, Volume 16, Issue 1, June 2014, Pages 1–10.
14. Zhou X, Qin X, Li K (2015) Parallel Techniques for Large Data Analysis in the New Version of a Futures Trading Evaluation Service. Big Data Research 2(3):102–109
15. Elastic (https://www.elastic.co/products/elasticsearch/)
16. Alluxio (http://www.tachyonproject.org)
17. ELK Stack (https://www.elastic.co/webinars/introduction-elk-stack)
18. Jin X, Wah BW, Cheng X, Wang Y (2015) Significance and Challenges of Big Data Research. Big Data Research 2(2):59–64
19. Krätzig M (2007) A software framework for data analysis. Comput Stat Data Anal 52(2):618–634
20. Mohammad Naimur Rahman (2016) and Amir Esmailpour, "A Hybrid Data Center Architecture for Big Data." Big Data Research 3:29–40
21. Pääkkönen P, Pakkala D (2015) Reference Architecture and Classification of Technologies, Products and Services for Big Data Systems. Big Data Research 2(4):166–186
22. Apache Kafta (http://kafka.apache.org)
23. Marathon (https://mesosphere.github.io/marathon/)
24. Apache Spark (http://spark.apache.org)

Mobile Application to Digitalize Education Volunteer Process

Aishwarya Chengalvala, Divya Sai Asanvitha Gundala, Nallamothu Lakshmi Hemanth, Sethu Sandeep Kadiyala, and G. Kranthi Kumar

Abstract Sharing knowledge has always been a complicated task from long ago, but the emergence of smartphones has made the task easier and more flexible. There are still many places that do not move with time, and hence, people can only learn limited knowledge with the limited workforce. Many educated people have come forward in improving these areas through various means such as NGOs and student volunteers. But there is no proper mode of communication between school management and the people who are interested in helping them. A mobile app is one possible solution to this problem. It connects schools, students, and student volunteers, making the communication between them simpler. Many features like slot booking, nearest school location details, chatting service, etc., are possible to incorporate inside the mobile app. This can be implemented by using Java programming language, Android Studio, and Firebase which is used to develop mobile applications. The final result is making these educational places have an efficient workforce, introducing newer ways to learn and education for everyone.

Keywords Mobile app · Android studio · Firebase · Student volunteers · Nearest schools · Slot booking · Education

1 Introduction

Android is an operating system developed for mobile devices, such as smartphones and tablets, based on the Linux operating system. It was developed by Google in 2005. It is a platform for smartphones.

Not all children are fortunate enough to have access to the internet, get trained by professionals, and get exposure to various domains of knowledge. There are hundreds of volunteers across the country who are willing to share their knowledge with young and thriving minds. On the other hand, volunteering helps one to learn different things

A. Chengalvala · D. S. A. Gundala · N. L. Hemanth · S. S. Kadiyala · G. Kranthi Kumar (✉)
Department of Computer Science and Engineering, V R Siddhartha Engineering College, Vijayawada, India
e-mail: kranthi@vrsiddhartha.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_5

about people and their cultures which can lead to one's growth. We can no doubt say that learning is both ways. The problem comes with how this can be made possible and how to bridge the gap between the two parties. Unfortunately, there is a high chance that volunteers are not aware of the children in need around them, also school management or students have no way to approach the interested volunteers. Nowadays, we have transforming technologies that are used in many various fields to solve problems like booking and reservations [1–3]. Evoking these technologies and merging them with a single platform to integrate students, management, and volunteers are our primary objective.

The volunteering platform is used to build a self-sustainable ecosystem for the growth of the underprivileged. Volunteers who invest their time will teach the subject(s) they understand well and will work closely with different stakeholders including principals, parents, and students [4].

2 Motivation

The vision of providing quality education to children in rural India. It is achieved by using technology to connect enthusiastic people from all over the world as volunteer teachers with public primary schools in remote villages in India. It focuses on children in the 8–12 age group (grades 3–7), to provide interactive classroom life in the local environment through a strong, relevant ecosystem.

To apply any of the above features at a school is not cost-effective, both management and volunteers prefer something handy and ready to use. Android is a software service designed for mobile devices, including key operating systems, middleware, and applications. Android apps are apps that run on the Android platform. Since the latter is designed for mobile devices, Android applications are designed for smartphones or tablets running on the Android operating system.

Therefore, a mobile application best suits our use case [5]. A case study of the most downloaded apps gives an understanding of designing an application with a practical user interface and a satisfying user experience [6]. It shows the real-time list of schools and allows you to book a slot in advance for the desired school, and finally you can navigate to the nearby school and teach the needy.

The ideology behind our implementation is as follows:

"EduHelp" application is used by both management and students to collaboratively work toward the functioning of the application.

The server-side processing will be enabled using Firebase [7, 8]. The client requests the server for locations, where the school is available, and the server responds with slot availability.

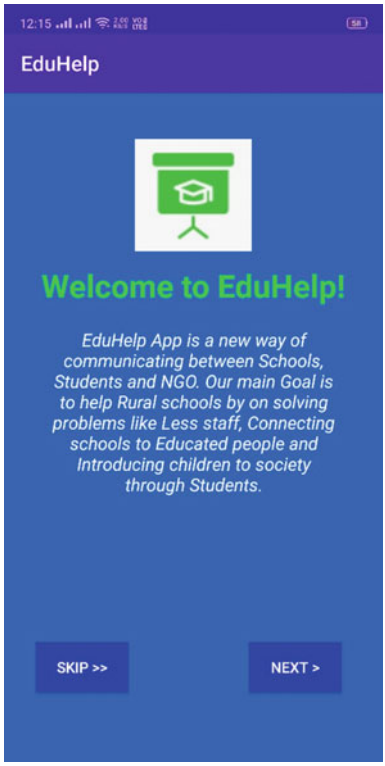


Fig. 1 Welcome screen of the app

3 Implementation

3.1 Starting the Application

The user has to install the application "EduHelp" on the Android-based device. After installation, the icon of the application will be displayed on the home screen of the user's device. "EduHelp" welcome screen (see Fig. 1) will be flashed to the user on opening the application followed by an optional instructions page (see Fig. 2).

3.2 Registration

When a user enters for the first time, they have to register to log in (see Fig. 3). This is a one-time registration. The user has to enter details like full name or the school name, phone number, email address, and password. If one is on behalf of the school management, then they must select "Register as a School". However, if one is a volunteer, then they must select "Register as a Student". All the data given will

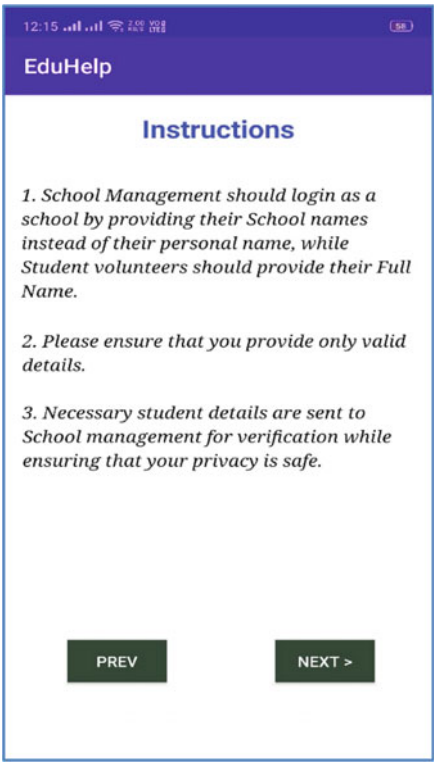


Fig. 2 Instructions page

be stored on the server, i.e., Firebase. Booking for slots is mandatory, and it has to be done before arrival.

3.3 Login

Once the user registers, he/she can use his/her email ID and password to login in the future. One can log in as school management or a volunteer (see Fig. 4). Based on their choice, authentication is done, and the client is sent to their respective pages.

3.4 Management View Screen

The school management will be provided with the details of the volunteers (see Fig. 5) who have booked their slots together with other details like date and the subject that will be taught (see Fig. 6).

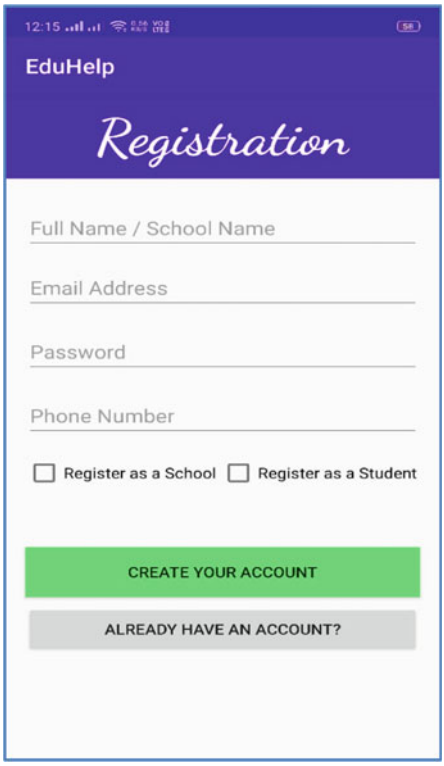


Fig. 3 Registration page

3.5 Selection of Location for Volunteering

The client is provided with multiple school locations. The client has to select one of the locations provided when he/she desires to teach the students. Later, the client will be provided with the name, time, date, and subject he wants to teach options (Fig. 7).

3.6 Confirmation

A confirmation message pops up on the screen for the volunteer after he books his slot for volunteering saying his slot is successfully booked. Email confirmation and SMS confirmation features can be included in the future.

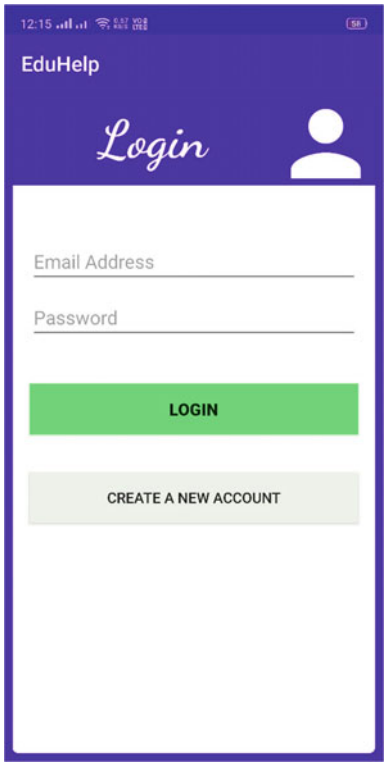


Fig. 4 Login page

4 Advantages

The EduHelp app made it possible to increase the teaching staff while helping children better connect to the world and providing them with quality education. Additionally, it helps bridge the gap between learners and explorers. Any educated person can teach in any school which shows the main purpose of the EduHelp app.

5 Limitations and Future Scope

The flexible communication between school management and student volunteers also makes it possible to simplify teaching in remote places. For now, we have incorporated only one school. However, more schools and volunteers across pan India shall be included subsequently. Some more functions that can make the app more robust are school location tracking, a better user interface, user verification, confirmation through email and SMS after booking, and control over the subjects that can be taught.

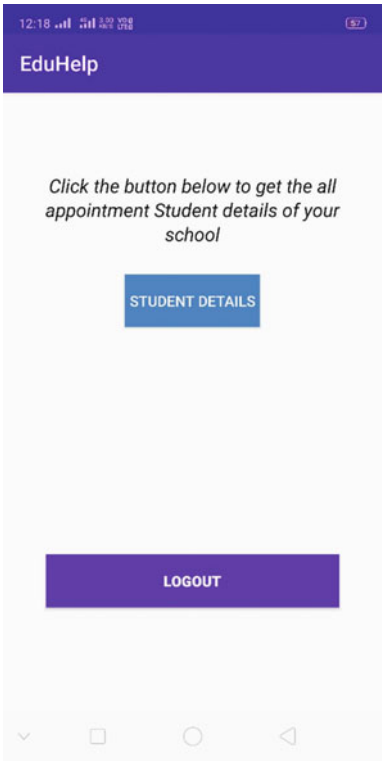


Fig. 5 Management view screen

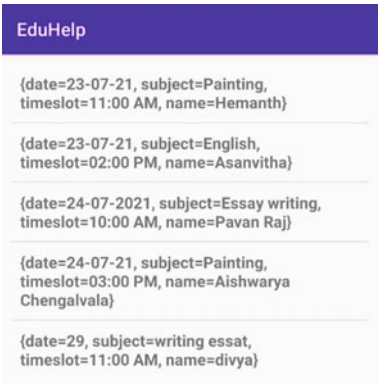


Fig. 6 Details of the booked slots

6 Conclusion

This paper summarizes an efficient way to connect students, schools, and volunteers using recent technology. This app allows the user to view and book slots for managing and teaching, respectively. Usage of this app at a large scale would benefit users albeit

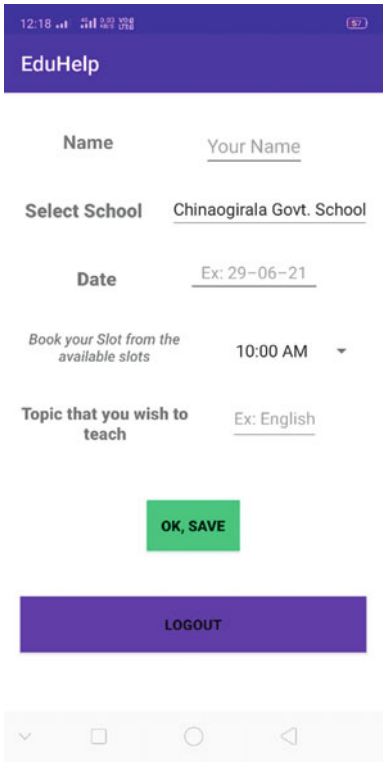


Fig. 7 Volunteer view screen

a user is in a new place. The app is user functional and convenient to use by people of all age groups.

References

1. Muddana K, Matsa D, Neerukonda K, Lahori HB, Kranthi Kumar G (2020) Android application for vehicle parking system parking informatics. IRJET 7:4841–4844
2. Jagtap SS, Hanchate DB (2017) Development of android based mobile app for prestashop eCommerce shopping cart (ALC). IRJET 4:2248–2254
3. Tenagyei Edwin K, Kwadwo K, Patamia Agbeshie R (2019) Design and implementation of hospital reservation system on android. IJCSIS 17:31–37
4. Hans G, Harsha S (2018) Mobile learning application and its usage among students in education. JETIR 5:984–998
5. Murugesan S, Deshpande Y (2002) Meeting the challenges of web application development: the web engineering approach. ICSE 24:687–688
6. Heonsik J (2017) A study on understanding of UI and UX, and understanding of design according to user interface change. IJAER 12:9931–9935
7. Sonam K, Swapnil T (2017) Real time databases for applications. IRJET 4:2078–2082
8. Jasra N, Owais Noor T (2018) Database for mobile application. IJTSRD 2:853–854

Emotion Recognition in Human Face Through Video Surveillance—A Survey of State-of-the-Art Approaches

Krishna Kant and D. B. Shah

Abstract Emotion is a state of thoughts and feelings related to events and happenings. Emotion plays a significant role in all aspects of life. If we capture the emotion of an individual, then many issues can be resolved. Emotion recognition is becoming an interesting field of research nowadays due to huge amount of information available from various communication platforms. Emotion detection will play vital role as we are moving towards digital era from all the aspects of life. The capacity to understand the emotion by computer is necessary in many applications, especially emotion detected from video. At the present, emotional factors are significant because we get efficient aspect of customer behaviour. Video surveillance plays important role in recent times for face detection and feature extraction. Video surveillance will help us to understand the emotion that is being carried out by an individual. Facial expression is the natural aspect which helps to integrate the quality of emotion. Human face helps to get the emotion a person expressing in due course of action and activities. Many clues we can get from the human face which will help us to formulate and solve the issues and threat in the domain of emotion detection. This scholarly work is a comparative study of emotion recognition and classification of emotion from video sequences which exhibits the current trends and challenges in emotion recognition. Intensity of emotion and duration of an emotion are the two significant attributes for effective identification of emotion of an individual though video surveillance. This study of emotion recognition has given insight and future direction in the domain of emotion recognition technology.

Keywords Emotion recognition · Human face detection · Video surveillance · Feature extraction · Emotion classification · Emotion intensity

K. Kant (✉)
Smt. Chandaben Mohanbhai Patel Institute of Computer Applications (CMPICA), Charusat University, Changa, Anand, Gujarat, India
e-mail: kantkrisna81@gmail.com

D. B. Shah
G.H Patel P G Department of Computer Science and Technology, Sardar Patel University, V.V. Nagar, Anand, Gujarat, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_6

1 Introduction

Emotion is a strong feeling which can be found when anybody makes decision about any situation, association, or happenings with others. It is necessary to recognize and anticipate human response in order to do estimation of one's emotional state in daily life. Various communication platforms can be used for detection of highly distrustful person from intelligent security which can match one's interest with video emotion; due to this reason, human being emotion has fascinated additional attention as a novel research trend. Video expression detection technique associates the video frames with human emotions for mining user emotion from present emotional stages and public affairs. Visual features comprise of light, dispersion, colour energy, colour temperature, film-transform rate, film period, film style change pace, motion strength and dynamics. To design a framework for video frame base emotion detection, we should have appropriate attribute frames of human face inside capacity. Video-based emotion detection is interdisciplinary which contains the field like psyche, perceptual compute and H.C.I. Research domain of psychology, nervous system has presented that human face is the utmost native main clue for collaborating the class and temperament of emotion which connects to human body and accent [1].

Emotion is crucial aspect which influences our daily life. Video is information carrying tool which has played a vital role in recent years in emotion recognition system [2, 3]. The human face contains of rich set of emotional indications. FACS was proposed that converts human face expression by a sequence of particular position gesture of appearance [4]. The AU can be recognized by geometric and facade attributes dig out from face image [5]. Video emotion detection develops the efficiency of video endorsement system. The psychological investigation is driven by user since it considers more mental foundation into consideration as we relate by conventional video investigation system [6, 7]. Videos uncover viewers to various types of optic and hearing vibrations that express story line knowledge and create various psychological experiences. It is significant to give emphasis to the association involving emotion and the attribute of video [8]. According to Gallup, the analytic and Advice Company, the report of optimistic emotions slightly reduced in 2017 compared with previous year. The survey also states that based on the report of positive emotions, the happiest country in 2019 was Paraguay has got the top ranking in consecutive three years. The least happy country was Afghanistan [9].

Mehrabia [10] defined that visual appearance of message contributes 55% on the whole feeling, whereas verbal element and syntactic information gives 38% and 7%, respectively. The basic emotion stated by Ekman's in 1970s were: sad, anger, disgust, fear, happiness and surprise. The entire paper is arranged as stated: Sect. 1 contains the introduction of Emotion Detection and Video processing, Sect. 2 contains related work, Sect. 3 contains the limitation of the study, Sect. 4 narrates results and discussion, Sect. 5 explains research gap, Sect. 6 includes proposed architecture, Sect. 7 states future direction, and Sect. 8 states conclusion.

2 Related Work

Number of researchers are working on emotion detection from facial expression; their works are stated as: Sikkandar proposed a face detection for better accuracy and efficiency for emotion detection. The used data set JAFFE, CK+, Pie data sets and few real-life pictures were used for their work. They have used DCNN for attribute extraction, ICSO for image recovery and ensemble the SVM and neural network to upgrade comprehensive correctness and recovery time of recommended framework. Their future plan is to develop mobile app for detecting emotion from each video frame. Their work outperforms existing system in terms of accuracy and decreased computation time [11]. Niu has established an algorithm with combination of oriented FAST and rotated BRIEF (ORB) feature and (LBP) for extracting model from facial expression to scale up the computational speed. For classifying the features extraction, they have used support vector machine and LBP. The used data set used in their study were JAFFE, CK+ and MMI. Authors have performed all the experiment on static images; in future, they will apply this algorithm in video sequence and will also focus on how to improve the speed and accuracy of the algorithm. Their result on JAFEE data set was 88.5%, on CK+ data set was 93.2% and on MMI was 79.85% [12].

Tian have studied personified emotion detection and forecasting model based on cloud computing consists of human face and emotion detection. Their forecasting model deploys SVM for face detection and deploy machine learning for emotion classification for understanding the motivation of identifying the emotional state through cloud computing. The result of the experimental depends on domain attribute achieves foremost and have good abstraction capacity and enhanced to 6.3% from conventional methods [13]. Arora developed automatic system for emotion detection. They have used JAFEE data set in their study. They proposed hybrid approach for feature extraction and optimization using principle component analysis (PCA) and particle swarm optimization (PSO). The PCA deployed for attribute mining and the PSO for optimizing the attribute vector as attaining good efficiency and accuracy in emotion detection. The author has deliberated classifiers such as decision tree, MLP and random forest tree. The classification rate of 94.97% was attained with the recommended method [14].

Gunawan illuminated on recognizing the emotions using convolution neural network and deep neural network techniques from the videos. This work was carried out on FER data set (2013) by applying HAAR Cascade method achieving the accuracy of 97% on training and 57.4% on testing data set. The researcher has also suggested that further intensity of scale will assist to project the potency of detected emotion. Multi-modal approach combined to design a model using various data sets [15]. Liu has advised a CAAN technique that is sturdy with respect to frames. They have also created a HEIV for better construct of fine video emotion detection framework. In future work, they will consider human body parts semantics for better human emotion recognition. Their CAAN attains 1.83% and 1.68 efficiency acquired on Ekman-6 and Video Emotion-8 [16].

Gupta presented a self-attention network that consists of CNN and ResNet. They achieved satisfactory result on FER data set in their study. Geometric feature-based method was applied to calculate the interval between various parts of the face for detecting emotion. Appearance-based method was applied for extracting feature vector for detection of emotion from the face. A model was proposed on FER data set attaining correctness 85.76% in training and 64.40% at testing phase [17]. Ivanova has devoted their work for developing emotion detection model which is centred on convolution neural network including the advanced method for creating training and testing specimen to establish the rational numbers of layers in neural networks. Viola–Jones method applied to resolve the obstacle of face detection in the study. Attains highest recognition accuracy of "happiness" 88%, the lowest of "sadness" 57%. During their experiment, a training data set of FER (2013) was applied for training and expression [18]. Zhang examined natural components in two-dimensional emotion chart through mining physical components of the video data such as colour, light intensity and sound. KNN, SVM and MLP classifiers of machine learning technique were used for accurately predicting emotion. ANOVA and principle component analysis were applied as classifier for statistics analysis. In future, they will start to collect more video clips to enhance accuracy [19].

El Dahshan established a method which comprises of deep belief network and quantum particle swarm optimization algorithm. They have used JAFFE and FER data set for their application. Feature extraction was carried out using sampling method and classification by the parameter of DBN which has been determined by the quantum particle swarm optimization during training phase. They achieved high accuracy and decrease the time of computation to 62% on JAFFE and 82% on FER-2013. In future, they have planned to merge ontology approach to add power to the system [20]. Afdhal developed an approach to find the classification of 12 emotion with their intensities. They have discussed novel emotion recognition methodology of primary and intensity emotions to address emotion states which are too subtle to discriminate. They use used deep learning for feature extraction and for classification they have used fuzzy logic. The author has made clear definition of emotion intensities of 12 emotions. They proposed a novel method for emotion recognition of 6 basic emotion and intensities of 12 emotions that has been identified. Their approach removes the misperception between the primary emotion and its intensities, which increases the rates of the error classification [21]. Mellouk has noticed that significant interest of researchers has increased using deep learning over recent years. They have discussed the different architecture of CNN and CNN-LSTM recently proposed by different researchers. Their work also describes the high rate of accuracy that has been achieved by different researchers in emotion detection. Author has motioned that not only FER provides six basic plus neutral emotion in day to day happenings. This motivates researchers to build new databases and create the different architectures to detect the all the basic emotion and secondary emotions. They have suggested to incorporate all the types of emotion that a human expresses in real life [22].

Hossain has proposed a edge cloud-based automatic face detection using IOT devices where they applied CNN and SVM for extraction and classification of emotion. They have compared their work with others on different data sets. They carried

out their experiment on RML and eNTERFACE'05 data set and achieved accuracy of 82.3% and 87.6%, respectively. In future, they will investigate other audio–video and cross-database; they also will study other deep fusion methods as an alternative of DSAE [23]. Vinola identified emotion intensity of happiness with average accuracy of 86.54 on AMFED data set. Geometrical features deliberated including balanced Euclidean distance of landmark points of eyes and lip part of human face. Fuzzy was applied to effectively categorize intensity of the emotion for video sequence. The major contributions of this paper were: (a) detecting smile, (b) constructing fuzzy rules controlled by the Euclidean distance multimedia tools and applications between the LMPs, (c) authorizing for slight movement of face in distinct frames of the video and (d) verification and justifying efficiency of various video frames in different scenarios [24]. Bahreini has proposed method for efficient emotion recognition and classification. The principle goal of work was to authenticate the usability of webcam data and accurate investigation of emotion in e-learning context. They have worked on CK+ data set for their experiment for emotion recognition. Their experiment presented that most intensive emotions as happiness, surprise detected better than less intensive emotions except neutral and fear. The proposed work converts facial aspects to recognized emotional states by the FURIA algorithm based on Fuzzy Logic. Their study attains correctness of 83.2%, equivalent to detection by human [25].

Adouani has introduced a relative techniques for face recognition from video sequences. The applied techniques were HAAR features, HOG with SVM and LBP. They have examined DEAP by applying physiological signals data set for their study. The proposed method was applied by Python with OpenCV and Dlib libraries. The outcome to the study identifies that the HOG and SVM technique is reliable and accurate than LBP and HAAR with detection rate of 92.68%. In future, they will consider working on the challenges to overcome such as induction variance, illumination, posture and face expression changed by testing other RECOLA, JAAFE and FRET databases. They also wish to implement their work using Artificial Intelligence methodology: fuzzy logic, ANN and deep learning [26].

Avots have introduced audiovisual information for emotion recognition. They used SAVEE, eNTERFACE'05 and RML and AFEW data set for training and testing process. Viola–Jones algorithm was used for face detection and facial image emotion classification was done by SVM and convolutional neural network (AlexNet). They attained the accuracy of 77.4% for SAVEE data set, 69.3% for RML, 50.2% for eNTERFACE'05 and 46.6% for AFEW. In future, they plan to add more database for their study and also they will incorporate to extend the data preparation where characteristics from single audio and facial image sample expressed in one vector [27]. Niu provided a comprehensive process to identify faces from video, which attains better efficiency in accuracy and speed. They applied Libfacedetection for face detection based on the integration of the LBP and AdaBoost technique. For feature extraction, dlib was used because of powerful detection capacity of deep learning. Gaussian mixed model (GMM) as a classification model is used for soft classification and predicts the similarities between the image of testing and training. The used data set was CAS-PEAL, FMedia and MaskNet [28]. Mo have used Hilbert–Huang

transform cross correlation (HHTC) and cross-correlation technique for emotional attribute in video detection method. Researchers have used support vector regression (SVR) a supervised machine learning algorithm and Ekman's emotion model to overcome the difference between low attribute and the perceptual states. Data set used were VACAD and the LIRIS-ACCEDE for effective video content analysis. They have suggested that constructed HHTC has good insight of operation on video than traditional features [8]. Wang has used support vector machine for facial recognition classification. For depression detection, they evaluate facial expressions and eye movement features by SVM. Comparative to number of depressed people, the number sample was less. They will explore more feature selection and classification techniques in future. The results of their research illustrate that these features can attain the detection accuracy rate at 78.85% and recall at 80.77% [29].

Baskar has studied machine learning algorithm for classification as: Bayesian network, hidden Markov model and SVM. They have also surveyed deep belief network, a machine learning method for face aspect detection. Hidden Markov model performs better classification for 3 or 5 state model by using self-determining or fusion of multimodel models. SVM algorithm presents better classification on all possible basic expression. Finally, deep belief network outperformed better and improved facial expression classification [30].

Giannakakis makes a study of a framework for recognition of stress and anxiety from recorded video. They achieved good accuracy from the facial clues obtained from eye, mouth, head movement activity and heart activity which distinguishes as a selective scale for stress and anxiety. For detection of features from the face, they applied active appearance models (AAM). The methods applied and verified were: KNN, generalized likelihood ratio, SVM, Naïve Bayes classifier and AdaBoost. The limitation of this study was the duration of the recorded video which is very less, it was suggested that it should be of at least one-minute duration [31].

Ithaya has presented a unique methodology for detecting facial expression with high recognition rate. They have applied Gabor wavelet method and multi-class AdaBoost for object detection and feature extraction and classification. Multi-class AdaBoost were used for selection and categorization for identifying the face expression. They have suggested that Gabor wavelet method outperforms better than HAAR features yielding higher detection and execution rate. Their recent study for identifying the facial emotion does not deliberate the fully automatic and occlusion not handled [32].

3 Limitations of the Work

Many researchers have given their contribution in the domain of emotion recognition based on facial expression from video data. Gunawan [15] in their study has not addressed the intensity of scale and multimodal approach for designing the model using various data set. Liu [16] has not included the human body parts semantics for better emotion detection. Abdulsalam [33] stated that the modalities like audio with

other data set have not been covered in their study. Zhang [19] has come with the fact that the collection of daily life video provides a better basis. Selvam [34] has not address the occlusion and its role in emotion classification. Salah [35] identified the gap in their study as: the annotation and ground truth acquisition of data, verification and validation for ground truth reality and lack of ecological validity in the annotations results. Giannakakis [31] found that duration of the recorded video is very less; it is suggested that it should be of at least one-minute duration. Selvam [34] states that they do not deliberate the fully automatic emotion detection, and occlusion has not been addressed. Hossain [23] applied CNN and SVM. Niu [28], Abdulsalam [33] have given the insight that the major aspect, i.e. intensity and duration in detecting the emotion without any subtle discrimination have not been addressed in the study. Afdhal [21] stated that emotion having intensity of 12 emotion on static images. Arora [14] has also has done their experiment on static images.

4 Results and Discussion

In this study, it has been identified that the importance of research in emotion recognition using video data has increased in current years. The video emotion recognition (VER) has certain steps to follow during the process like: pre-processing, face detection, design of model architecture and, finally, emotion recognition and classification. Pre-processing is the significant step in emotion recognition which was present in all the research paper studied in the comparative study. Several techniques for pre-processing, face detection,emotion recognition and classification used in the research work were: DCN and SVM were applied in their study [11], PCA and PSO for their experiment [14], deep belief network and quantum particle swarm optimization [20], context-aware attention network (CAAN) [16], LBP and SVM [12] CNN and DNN network [15], self-attention network deep learning: CNN and ResNet [17], convolution neural network and Viola–Jones [18], convolution neural network and extreme learning machine for feature mining and SVM [23], context-aware attention network and convolution neural network [16], fuzzy logic [24], Viola–Jones and deep convolution neural network [21], FURIA [25], dynamic reinforced local binary pattern (DRLBP, multi-class AdaBoost [34], deep neural network through transfer learning, ELM and LSTM [35], Gabor wavelet method and multi-class AdaBoost for emotion classification [32], PCA, ANOVA for aspect mining, and for classification, they have applied KNN, SVM and MLP [20] (Table 1).

5 Research Gap

From literature reviewed, it is identified that still there is scope to design and develop emotion detection of facial expression from video sequence. All the research work studied mainly focussed on the Ekman's emotion model in the study. The research gap identified during the review of this work are stated are as:

Table 1 Comparison of methods with results

Author name	Data set	Methods used	Results
Sikkandar [11]	CK+, Pie	DCNN, SVM	Increased accuracy and decreased computation time
Arora [14]	JAFFE	PCA and Particle swarm optimization	Rate of classification: 94.97%
Niu [12]	JAFFE, CK+, MMI	LBP, SVM	JAFFE: 88.5%, CK+: 93.2% and MMI: 79.85%
Gunawan [15]	FER	CNN, DNN	Training set: 97%, Testing set: 57.4%
Gupta [17]	FER	Self-attention network deep learning: CNN	Training set: 85.76%, testing set: 64.40%
Ivanova [18]	FER	Convolution neural network, Viola–Jones	Happiness 88%, sadness 57%
Bahreini [25]	CK+	FURIA	Accuracy: 83.2%
Hossain [23]	RML, eNTERFACE'05	CNN, SVM	RML: 82.3%, eNTERFACE: 87.6%
Vinola [24]	AMFED	Fuzzy logic	Average accuracy: 86.54%
El Dahshan [20]	JAFFE and FER	Deep belief network and quantum particle swarm optimization	JAFFE: 62%, FER: 82%
Liu [16]	HEIV	Context-aware attention network	1.83 and 1.68% performance gain on Ekman-6 model
Avots [27]	SAVEE, eNTERFACE'05, RML, AFEW	SVM, Viola–Jones, convolutional neural network	SAVEE: 77.4%, RML: 69.3%, eNTERFACE'05: 50.2%, AFEW: 46.6%

- Item Human body parts semantics [16]
- Audio–video modality [23]
- Occlusion and its role in emotion classification [34]
- Does not deliberate the fully automatic emotion detection and occlusion [32]
- Intensity of scale and multi-modal approach for emotion [15]
- Audio modality in emotion detection [33]
- The notation and attainment of data, magnifies condition for the estimation prototype that depends on professional annotation for depth reality, absence of environmental soundness in the notations [35]
- Video duration is less than one minute [31]
- Intensity not addressed for other basic types of emotion except surprise [24]

- According to the study made through several several research contribution in the field of emotion detection, intensity and duration of emotion has not been much explored in the study so far stated.

6 Proposed Architecture of Emotion Detection

The generic architecture of emotion detection comprises of five phases as: pre-processing, feature extraction, classification, intensity calculation and Happy Index. In pre-processing step, the video will be taken as input, then frames will be generated; after that, face will be detected from detected frame, then region of interest will be taken into consideration for the next phase. In the feature extraction, the different attributes from the selected region of interest will be taken into account. Classification phase will classify the types of emotion detected from the regions. The intensity calculation will calculate the intensity of six detected emotions. The last phase will generate Happy Index while performing operation on intensity of each detected emotion. The outcome of this architecture can be applied in various domain for better understanding of emotion of an individual.

7 Future Direction

Many researchers have given their contribution in the field of emotion detection which is very popular domain of research. The prime objective of this work is to discover the gap in emotion detection through video data. With the gap that we have identified, we will design a model which will bridge the gap and come out with a solution yielding efficient model. The feature extraction and classification will be carried out using novel methodology while considering the intensity of all the basic types of emotion. After obtaining the emotion with their intensity, the model can generate Happy Index of an individual on the basis of intensity of all the emotions. Video surveillance has become a weapon to get the insight from different facial expression of humans. The intensity plays a vital role in identification of accurate emotion classification which helps to reduce the classification error and enhancing the actual emotion identification.

8 Conclusion

The prime objective of the research paper is to compare various emotion recognition methods which are based on video input. Versatile video inputs have been obtained and processed by various researchers, and they have worked over these inputs through different machine learning and deep learning techniques. This paper includes the

study of the various methods from 2016 to 2021 with accuracy and pitfall and also presented a survey on facial expression recognition with current feature extraction techniques. The objective of this study is to identify the work that has been contributed in this domain so as to get the pitfalls which will help to understand the work to be carried out. This study has identified the research gap in the domain of emotion detection through video data. The study says that there is need to design a model in order to scale up the accuracy and efficiency of emotion recognition in terms of intensity and duration which will address the classification rate of emotion.

References

1. Cohn JF (2010) Advances in behavioral science using automated facial image analysis and synthesis [social sciences]. IEEE Signal Process Mag 27(6):128–133
2. Smith GM (2003) Film structure and the emotion system. Cambridge University Press
3. Dorai C, Venkatesh S (2001) Computational media aesthetics: finding meaning beautiful. IEEE Multimed 8:10–12
4. Chu W-S, De la Torre F, Cohn JF (2016) Selective transfer machine for personalized facial expression analysis. IEEE Trans Pattern Anal Mach Intell 39(3):529–545
5. Eleftheriadis S, Rudovic O, Pantic M (2014) Discriminative shared Gaussian processes for multiview and view-invariant facial expression recognition. IEEE Trans Image Process 24(1):189–204
6. Hanjalic A (2005) Adaptive extraction of highlights from a sport video based on excitement modeling. IEEE Trans Multimed 7(6):1114–1122
7. Hanjalic A (2006) Extracting moods from pictures and sounds: towards truly personalized tv. IEEE Signal Process Mag 23(2):90–100
8. Mo S, Niu J, Su Y, Das SK (2018) A novel feature set for video emotion recognition. Neurocomputing 291:11–20
9. Alshawatreya A, Malkawi M, AlBashiti A (2017) The detection of security threat level by analyzing certain human emotions using neuro-fuzzy techniques. In: 2017 intelligent systems conference (IntelliSys), pp 1122–1129. https://doi.org/10.1109/IntelliSys.2017.8324270
10. Scherer KR (1985) Handbook of methods in nonverbal behavior research. Cambridge University Press
11. Sikkandar H, Thiyagarajan R (2021) Deep learning based facial expression recognition using improved cat swarm optimization. J Ambient Intell Humaniz Comput 12(2):3037–3053
12. Niu B, Gao Z, Guo B (2021) Facial expression recognition with LBP and ORB features. Comput Intell Neurosci
13. Tian W (2021) Personalized emotion recognition and emotion prediction system based on cloud computing. Math Probl Eng
14. Arora M, Kumar M (2021) AutoFER: PCA and PSO based automatic facial emotion recognition. Multimed Tools Appl 80(2):3039–3049
15. Gunawan TS, Ashraf A, Riza BS, Haryanto EV, Rosnelly R, Kartiwi M, Janin Z (2020) Development of video-based emotion recognition using deep learning with Google Colab. TELKOMNIKA 18(5):2463–2471
16. Liu X, Wang M (2020) Context-aware attention network for human emotion recognition in video. Adv Multimed
17. Gupta A, Arunachalam S, Balakrishnan R (2020) Deep self-attention network for facial emotion recognition. Procedia Comput Sci 171:1527–1534
18. Ivanova E, Borzunov G (2020) Optimization of machine learning algorithm of emotion recognition in terms of human facial expressions. Procedia Comput Sci 169:244–248

19. Zhang J, Wen X, Whang M (2020) Recognition of emotion according to the physical elements of the video. Sensors 20(3):649
20. El Dahshan KA, Elsayed EK, Aboshoha A, Ebeid EA (2020) Recognition of facial emotions relying on deep belief networks and quantum particle swarm optimization. Int J Intell Eng Syst 13:90–101
21. Afdhal R, Ejbali R, Zaied M (2020) Primary emotions and recognition of their intensities. Comput J
22. Mellouk W, Handouzi W (2020) Facial emotion recognition using deep learning: review and insights. Procedia Comput Sci 175:689–694
23. Hossain MS, Muhammad G (2019) Emotion recognition using secure edge and cloud computing. Inf Sci 504:589–601
24. Vinola C et al (2019) Smile intensity recognition in real time videos: fuzzy system approach. Multimed Tools Appl 78(11):15033–15052
25. Bahreini K, Van der Vegt W, Westera W (2019) A fuzzy logic approach to reliable real-time recognition of facial emotions. Multimed Tools Appl 78(14):18943–18966
26. Adouani A, Henia WMB, Lachiri Z (2019) Comparison of Haar-like, HOG and LBP approaches for face detection in video sequences. In: 2019 16th international multi-conference on systems, signals & devices (SSD). IEEE, pp 266–271
27. Avots E, Sapiński T, Bachmann M, Kamińska D (2019) Audiovisual emotion recognition in wild. Mach Vis Appl 30(5):975–985
28. Niu G, Chen Q (2018) Learning an video frame-based face detection system for security fields. J Vis Commun Image Represent 55:457–463
29. Wang Q, Yang H, Yu Y (2018) Facial expression video analysis for depression detection in Chinese patients. J Vis Commun Image Represent 57:228–233
30. Baskar A, Kumar TG (2018) Facial expression classification using machine learning approach: a review. Data Eng Intell Comput 337–345
31. Giannakakis G, Pediaditis M, Manousos D, Kazantzaki E, Chiarugi F, Simos PG, Marias K, Tsiknakis M (2017) Stress and anxiety detection using facial cues from videos. Biomed Signal Process Control 31:89–101
32. Rani PI, Muneeswaran K (2017) Recognize the facial emotion in video sequences using eye and mouth temporal gabor features. Multimed Tools Appl 76(7):10017–10040
33. Abdulsalam WH, Alhamdani RS, Abdullah MN (2019) Facial emotion recognition from videos using deep convolutional neural networks. Int J Mach Learn Comput 9(1):14–19
34. Selvam IRP, Prasath TH (2019) Dynamic eyes and mouth reinforced LBP histogram descriptors based on emotion classification in video sequences. In: Smart techniques for a smarter planet. Springer, pp 185–209
35. Salah AA, Kaya H, Gürpınar F (2019) Video-based emotion recognition in the wild. In: Multimodal behavior analysis in the wild. Elsevier, pp 369–386

Character Segmentation from Offline Handwritten Gujarati Script Documents

Mit Savani, Dhrumil Vadera, Krishn Limbachiya, and Ankit Sharma

Abstract Modern innovations make great impacts on the human lifestyle and their way of working. It boosts the efficiency and productivity of people by reducing efforts, which help to handle several tasks at a time. Nowadays, all the government offices, banks, businesses, and education systems are influenced by paperless technology. It improves documentation and the secure sharing of information by saving space and resources. Paperless technology works on Optical Character Recognition (OCR) to convert all physical documents into machine-based documents. OCR consists of mainly two steps: segmentation, and recognition. The success rate of character recognition depends on the segmentation of required regions of interest. This paper introduced an algorithm that utilized projection profile, bounding box, and connected component labeling techniques for the development of Gujarati handwritten dataset and segmentation of handwritten text from Gujarati documents into line, word, and characters.

Keywords Handwritten document segmentation · Bounding boxes · Connected component labeling · Projection profile technique

M. Savani (✉) · D. Vadera · K. Limbachiya · A. Sharma
Department of Instrumentation and Control Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India
e-mail: 17bic046@nirmauni.ac.in

D. Vadera
e-mail: 17bic060@nirmauni.ac.in

K. Limbachiya
e-mail: 20ftphde43@nirmauni.ac.in

A. Sharma
e-mail: ankit.sharma@nirmauni.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_7

1 Introduction

People working at educational institutes, government officials, and private businesses having many difficulties dealing with physical documents. Finding particular details from the paper-based document and updating those details are tough and time-consuming tasks. Even these physical documents require large space to store them appropriately. It is a wastage of time and resources like paper and money. Optical Character Recognition (OCR) is a solution for this problem, which helps to store, read, and modify data easily. The OCR system converts all physical documents into machine-based documents. Users can easily read, write; update the details from machine-based documents in a short time. Optical Character Recognition-based natural language processing analyzes documents by utilizing several steps such as preprocessing, segmentation, feature extraction, recognition, and postprocessing to process documents.

Processing of machine-printed documents is easier than handwritten documents because of uniformity in characters' spacing and writing style and size. Irregularities in handwritten characters create a big issue of false segmentation, which reduces the recognition accuracy of the overall system. Segmentation of the region of interest can boost the performance of the system. Many segmentation methods are popular, but not so efficient to find a correct region of interest. Here, we modified several parameters from popular techniques to get a region of interest accurately to achieve a good recognition rate.

2 Literature Survey

The earlier work for Gujarati OCR worked on printed Gujarati documents, which is quite easy due to the uniform size of characters and proper spacing between them. Working with a handwritten document is difficult due to handwriting variation. Handwriting presents skew variation, size variation, and incorrect spacing between lines and characters. Here, we provided a comprehensive survey for different segmentation techniques proposed for different languages.

Rajyagor and Rakholia [1] suggested a segmentation methodology that considers connecting characters, line slop, character overlapping, and other factors. Line segmentation was done using the horizontal projection approach, word segmentation was done using the scale-space technique, and character segmentation was done using the vertical projection strategy. Over a sample of 500 + photos, they attained an accuracy of 82% for character segmentation, 90% for word segmentation, and 87% for line segmentation. According to Tamhankar et al. [2], the separation of individual characters is the most challenging module. They believed that the rate of character recognition is highly influenced by character segmentation accuracy. To reduce the character segmentation error on the MODI script, which has no separation

between two consecutive words, they used a dual thresholding criterion. The algorithm was evaluated on over 15,000 characters, with over 10,000 of them correctly segmented, yielding 67% accuracy. Dahake and Sharma [3] explained an approach to deal with word segmentation from online handwritten Gurmukhi statements. For the separation of words from a sentence, the thresholding method is utilized. The vertical gaps between the strokes are first identified, and then the word is retrieved from the text using the maximum threshold value. The proposed method was put to the test on 200 sentences. With the use of the algorithm, an accuracy of 91.0% was achieved in segmentation. Jindal and Jindal [4] proposed the midpoint detection-based approach for segmenting lines and words from the handwritten Gurumukhi script, which features skewed lines, overlapped lines, and connected components between adjoining lines. The approach relies on the detection of gaps between two lines or words. The accuracy was found to be 95% in the line and word segmentation. Saha and Bal [5] introduced a technique using modified horizontal and vertical projection for segmentation of line and word from a document having multiple skewed text lines and overlapping of characters. A baseline recognition and normalization approach based on orthogonal projections was combined with a pressure detection method to determine a writer's personality. The method was tested over more than 550 text images of the IAM database and obtained segmentation accuracy of 95.65% over lines and 92.56% over words. With a very low error rate, the technique correctly normalizes 96% of lines and words. James et al. [6] demonstrated a unique approach for segmentation of overlapped, non-uniformly skewed, and touching text lines. For text line extraction, the method employed a piece-wise water flow methodology. The Spiral Run Length Smearing Algorithm is used to separate words from lines. Each word's skew is determined, and skew correction is done at the word level. The Connected Components Labeling method is used to extract the constituent characters. In Malayalam handwritten character recognition, the technique outperforms previous methods. Method of segmentation of Devnagri characters was proposed by Sahu and Sahai [7] using a bounding box, rectangle function, and region props operation on MATLAB. They used Otsu's threshold for the conversion of an image into binary format. The authors proposed a segmentation approach for line, word, and characters and achieved 100% accuracy for line and word segmentation, where 90%accuracy for character segmentation. Tan et al. [8] described a nonlinear clustering-based handwritten character segmentation technique. The technique divides a text line into strokes and computes a similarity matrix based on the stroke gravities. The cluster labels for these strokes are obtained using nonlinear clustering techniques. They utilized two nonlinear clustering methods: segNcut, which is spectral clustering based on Normalized Cut (Ncut), and segCOLL, which is kernel clustering based on Conscience On-Line Learning (COLL). They applied the SegNcut method on IAM, KAISTHanja1, HCL2000, and HIT-MW and achieved a correct rate of 85.4%, 89.2%, 88.9%, 89.7%. They applied another approach called segCOLL to all those databases and achieved accuracies of 88.1%, 85.2%, 89.3%, 87.7%. Arefin et al. [9] suggested the distance-based segmentation (DBS) approach for segmentation problems in lines, words, and characters in Bangla handwritten character recognition. To eliminate the shadows from the input image, they used the adaptive thresholding

approach. They utilized segmented characters as ROI to extract HOG features and sent them to SVM for classification. The technique achieved an accuracy of 94.32% for character recognition but it is incapable to segment joint and italic characters. Alaei et al. [10] developed a unique technique for unconstrained handwritten text segmentation into lines. They utilized a painting method that improves the separation of the foreground and background sections, making text lines easier to identify. Applied morphological dilation and thinning operations subsequently with some trimming operations to obtain several separating lines. They analyzed the distances between the starting and ending points by using these separating lines. The suggested technique was tested upon documents in English, German, Greek, French, Persian, Bangla, and Oriya with remarkable results. The existence of the header line, overlapping of characters in the middle zone, touching characters, and conjunct characters complicate the segmentation process, according to Gurjar, Deshmukh, and Ramteke. They experimented with segmenting handwritten Marathi text documents into lines, words, and characters depending on content and non-content. The results showed an accuracy of 83%, 90%, and 86% for the segmentation of lines, words, and characters. The proposed approach only works for isolated characters and not for touched characters [11]. Swetha et al. [12] have proposed two algorithms the Vertical Strips and the Connected Components techniques over a database consisting of approx 200 images to evaluate the performance of algorithms. The dataset is classified into three categories: easy, medium, and complicated. They showed an overall line recognition rate on three different types of images, each with a different level of complexity. For the easy and medium categories of photos, the Connected Components approach outperforms the Vertical Strips method. For a Complex category of images, the Vertical Strips method outperforms the Connected Components method. The algorithm is limited to data having uniform line spacing and word size.

3 Proposed Method for Segmentation

The method introduced in this paper is used for two aspects such as dataset generation for Gujarati language and extraction of text in terms of characters from handwritten Gujarati form. As we discussed that specific part of an image that we can call "Region of Interest" is quite important to obtain the best quality for any object recognition or identification-based applications.

3.1 Dataset Generation

Here, we used handwritten forms filled by different people in different writing styles for dataset generation as shown in (see Fig. 1). Due to the presence of grids, the segmentation task was difficult by normal Contour Formation techniques. Even when these grids are removed, the quality of Characters is degraded. We performed several

Fig. 1 Dataset form

logics using different techniques like Connected Components, Hough Transforms, or Kernel Generation for data segmentation. Not all those were enough to give the results we expected. Finally, we tried a fusion approach of projection profiles with vertical and horizontal profiles and bounding boxes to form contours and to extract the data inside of them. We applied vertical and horizontal profiles separately on forms and combined their result to get grid line contours.

Our algorithm for dataset generation works as per the following steps:

- Scan the image into binary format to make it ready for image preprocessing (see Fig. 1).
- Apply preprocessing techniques such as canny edge detection to detect edges and Gaussian Blur filter to reduce noise (see Fig. 2).
- Generate horizontal and vertical kernels separately with minimum width to detect horizontal and vertical lines by neglecting characters present in forms.
- Apply opening operation of image morphology to remove noise from detected straight horizontal and vertical lines.
- Addition of results from horizontal and vertical kernels to obtain an image in form of grids (see Fig. 3).
- Find an area of contours using specific limits and draw bounding boxes around the characters.
- Sort all detected contours from top to bottom vertically and left to right horizontally.
- Use the pixel address of each bounding box to cut out the area of interest with the best quality from the test image.
- Store cut-outs of characters in separate folders according to class (see Fig. 4).

Fig. 2 Form preprocessing

Fig. 3 Contour detection

3.2 Character Segmentation of Gujarati Document

Segmentation of characters from a handwritten document can be done in the same way as the dataset generation procedure. Here, we scanned a handwritten Gujarati text document and converted it into a digital image (see Fig. 5). Digital image of text

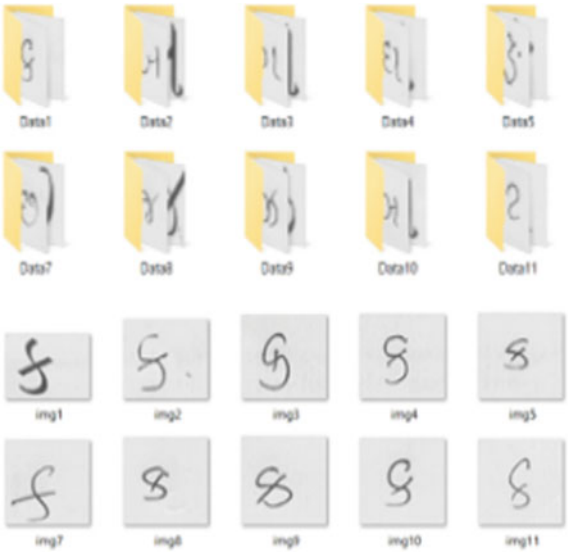


Fig. 4 Store image cut-outs

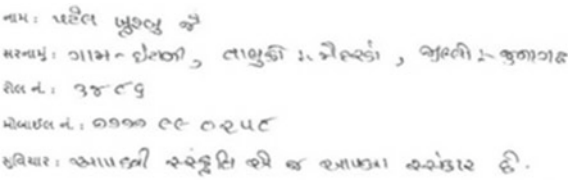


Fig. 5 Input scanned document

document was processed through the OpenCV library to convert into a grayscale image, to correct the skew angle of the image, and to apply thresholding operation. Thresholding converts pixel values above a certain threshold into white (i.e., 255 value) and the rest are converted into black (i.e., 0 value). If vertical lines are present in the form image, it can generate noise.

Morphological opening operation with vertical kernel applied to remove noise generated through vertical lines. Finally, we obtained a black and white image in which the bright regions are the handwritten or printed words or characters considered as foreground images, and the dark region in space is considered as background in the pre-processed image (see Fig. 6).

Fig. 6 Dilated image of a scanned document

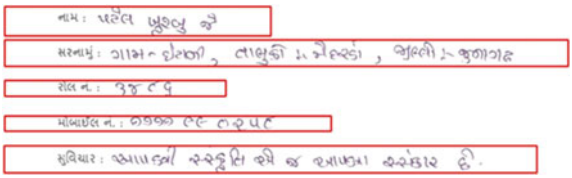


Fig. 7 Detected lines from the document

Fig. 8 Dilated image of a single line

Fig. 9 Detected words in a single line

We utilized several functions available in the OpenCV library such as Find Contour to extract the contours from the binary image, RETR_TREE, and CHAIN_APPROX_SIMPLE as contour approximation methods to compress and remove all redundant points from the contour. In this way, we segregated the lines from handwritten document images (see Fig. 7).

The process of word segmentation from detected lines was done as similar as line segmentation but we used a loop over the segmented lines sequentially (see Figs. 8 and 9). Applied resize operation on segmented word images to maintain the accuracy of the algorithm.

Finally, characters were segmented using the discussed method for line and word segmentation. We applied gray scaling, binarization, thresholding, dilation, contours detection along the bright region, and obtained the individual characters from handwritten forms (see Figs. 10 and 11).

Fig. 10 Processed word image

Fig. 11 Characters segmentation

4 Result and Discussion

The proposed algorithm can be used in areas of Banking, Education, Defense, Information Technology, etc. as a solution for Gujarati Character Segmentation. The algorithm can be clubbed with OCR models and various Digital formatting tools that convert the Physical Documents to Digital Format. In addition, the algorithm can be used to develop a Dataset for different languages. Here, we processed an approach for image segmentation that is useful for dataset generation as well as document and form segmentation. We utilized several functions from the OpenCV library to develop it. We tried our algorithm for segmentation of 50 grid forms (see Fig. 1) to generate a dataset for the Gujarati language and the results are as expected and satisfactory. Once we tried this algorithm for 50 images of handwritten Gujarati documents (see Fig. 5) for segmentation, it worked efficiently for characters with uniform sizes presented in a horizontal line. Sometimes it caused an error to detect details available in slant lines and segment a few portions of characters because of discontinuity in characters' shape.

5 Conclusion

Handwritten forms have great complexity because of variation in handwriting from person to person, variation in the use of pencil/pen with different ink colors, and text width variation. The proposed method is used for two different tasks such as dataset generation from handwritten or machine-printed grid-based forms and segmentation of handwritten or machine-printed details from actual forms available at schools, colleges, banks, etc. We used a cascaded way, where outputs of line segmentation are provided to word segmentation and followed by character segmentation to extract essential details. In the proposed algorithm, we used the projection profile technique with vertical and horizontal profiles and the bounding box method to get the region of interest. Overall the method worked accurately to segment details for stated purposes. This segmentation work can be applied to the development of the Gujarati language OCR system in the future to boost performance.

References

1. Rajyagor B, Rakholia R (2021) Tri-level handwritten text segmentation techniques for Gujarati language. Indian J Sci Technol 14(7):618–627. https://doi.org/10.17485/ijst/v14i7.2146
2. Tamhankar PA, Masalkar KD, Kolhe SR (2020) A novel approach for character segmentation of offline handwritten Marathi documents written in MODI script. Procedia Comput Sci 171(2019):179–187. https://doi.org/10.1016/j.procs.2020.04.019

3. Dahake D, Sharma RK, Singh H (2017) On segmentation of words from online handwritten Gurmukhi sentences. In: Proceedings of the 2017 2nd international conference on man and machine interfacing (MAMI) 2017, vol 2018, pp 1–6. https://doi.org/10.1109/MAMI.2017.8307870
4. Jindal P, Jindal B (2016) Line and word segmentation of handwritten text documents written in Gurmukhi Script using mid point detection technique. In: 2015 2nd international conference on recent advances in engineering and computational sciences (RAECS) 2015. https://doi.org/10.1109/RAECS.2015.7453388
5. Bal A, Saha R (2016) An improved method for handwritten document analysis using segmentation, baseline recognition and writing pressure detection. Procedia Comput Sci 93:403–415. https://doi.org/10.1016/j.procs.2016.07.227
6. Haji SAB, James A, Chandran S (2016) A novel segmentation and skew correction approach for handwritten Malayalam documents. Procedia Technol 24:1341–1348. https://doi.org/10.1016/j.protcy.2016.05.140
7. Sahu N, Sahai M (2016) Segmentation of Devanagari character. In: 2016 3rd international conference on computing for sustainable global development (INDIACom), 2016, pp 4005–4008
8. Tan J, Lai JH, Wang CD, Wang WX, Zuo XX (2012) A new handwritten character segmentation method based on nonlinear clustering. Neurocomputing 89:213–219. https://doi.org/10.1016/j.neucom.2012.02.026
9. Arefin N, Hassan M, Khaliluzzaman M, Chowdhury SA (2018) Bangla handwritten characters recognition by using distance-based segmentation and histogram oriented gradients. In: 5th IEEE region 10 humanitarian technology conference (R10-HTC), vol 2018, pp 678–681, 2018. https://doi.org/10.1109/R10-HTC.2017.8289049
10. Alaei A, Pal U, Nagabhushan P (2011) A new scheme for unconstrained handwritten text-line segmentation. Pattern Recognit 44(4):917–928. https://doi.org/10.1016/j.patcog.2010.10.014
11. Ramteke S, Gurjar AA, Deshmukh DS (2016) Automatic segmentation of content and noncontent based handwritten Marathi text document. In: Proceedings of the international conference on global trends in signal processing, information computing and communication ICGTSPICC 2016, pp 404–408. https://doi.org/10.1109/ICGTSPICC.2016.7955335
12. Swetha S, Chinmayi P, Mamatha H (2019) Line segmentation of handwritten Kannada documents. In: 2019 10th international conference on computing, communication and networking technologies (ICCCNT) (2019):1–5

A Comparative Study of Stylometric Characteristics in Authorship Attribution

Urmila Mahor and Aarti Kumar

Abstract This paper presents a study of various stylometric characteristics and their uses along with various methods of feature selection, extraction, classification for authorship attribution (AA). We learned of the importance of data or corpus size, which plays a heuristic role in the choosing of a sufficient amount of features to appropriately recognize the true writer of an undesignated text or piece of document. We also analyzed that by the time features type also changed and new features introduced that made this task easier. Over time machine learning became more interactive and easy to use. Many S/W are available for the AA task like WEKA, JAGAAP.

Keywords Authorship attribution · Stylometry · Corpus size · Feature set · Classification

1 Introduction

Many people write documents in multiple ways. Each person's writing style depends on mood, topic, and language. A person can express his/her idea using the different words in different moods so we can say that a single sentence can be written by the same person differently but if a person writes anonymously then he/she uses some words frequently/habitually unconsciously. AA helps us to know the genuine or actual or original writer of a certain document by identifying the author's writing style or technique. AA is needed when two persons claim their ownership for a single/the same document. Now a day's AA is the latest field of forensic. In AA, many attributes have been introduced as feature characteristics like lexical, syntactic features, structured, and content specific. AA is an analysis leading discover the

U. Mahor (✉) · A. Kumar
Rabindra Nath Tagore University, Bhopal, Madhya Pradesh, India
e-mail: urmila.mahor@gmail.com

A. Kumar
e-mail: aartikumar@aisect.org

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_8

original authorship [1–3] used word length distribution for recognizing writer's style writing [4] mentioned in his work, the plenty of vocabulary type features in his studies. Automated or self-regulating authorship analysis [5–7] received increasing attention in CS/IT for document reorganization with encouraging practical applications in several areas, like plagiarism detection or forensic sciences [8]. In AA research, writing characteristics and style characteristics are more advantageous [9, 10], as contemporaneous, features related to content are stronger in documents written by one person or the same person. Nowadays, most researchers insist on effectively, efficiently extracting style features from documents which do not influence by the particular content or writing style [11, 12]. Feature/characteristics set can be partitioned into the following categories.

2 Basic Feature Set

In this set lexical, syntactic, structural, and content specific characteristics are encountered.

(a) **Lexical**

Many researchers have used lexical features in their studies with various classifiers and methods. *N*-grams apply to so many languages and any corpus with tokens is the main advantage of this feature. Lexical *n*-grams are also another lexical feature that is popular and most useful as of their effectiveness if we compare it with *n*-gram, whether they are character or syntactic *n*-grams, although all possible *n*-gram features [13].

Characters and words *n*-grams are famous and known as one of the useful features for the work of stylometry for small size texts. Character *n*-gram AA [14, 15] and can be easily measured in so many languages without any specialization. An amalgam of features like word m-gram, character level *n*-grams gives good and accurate results because they are complementary to each other [3]. Character *n*-gram features are robust and performed well even the training corpus and testing data were taken from topics different [16]. In recent studies, we have noted that *n*-grams, function words, bag of words, and part-of-speech have emerged as reasonable substitutes for style features. Character *n*-gram is more dominant [17, 18]. Some researchers used sentence length and character *n*-gram [15, 19–21], [163]. Character *n*-grams can be good predictors as they have been recognized as successful features in many classification tasks of AA [15]. Replacement of function words (FW) are character *n*-gram was introduced by [17].

Function words are useful in resisting the stylistic imitation or copy. Some also take *n*-grams as the stem of the word. Research shows that text document characterization based on FW is used in languages as Word Stem supplement with information from the low-frequency level [10, 12]. This helps in explaining that these types of features prove more robust on cross-domain/genre problems [16, 18]. Function word is also a new feature set, first research using function word for AA was conducted by Burrows

(1987) [22], and this research used articles, prepositions ad pronouns as a feature vector. Function words are the most widely used lexical features and play important role in capturing the differences in writing styles [23]. Typically, the frequency values of function words are utilized to represent features. One of the extensively used features is FW for AA task.

(b) **Syntactic features**

These are generally severed as deep linguistic characteristics. It is cumbersome to manipulate intentionally. Typically, part-of-speech [POS] tags and *n*-grams features are utilized for authorship. Syntactic information is represented in the grammatical sequence of sentences [24, 25]. Syntactic characteristics have been regarded as more authenticate style attributes than lexical features because syntactic features are not controlled by the writer and cannot be easily manipulated [19, 26, 27]. POS information is more effective and can reflect reliable authorial fingerprint than lexical information [28]. Syntactic feature sets are faithful predictors for authoritative attribution [29]. The major advantage of the syntactic feature and FW is, they cannot be easily altered by the writer, so they give ample hint in attribution [30, 31].

(iii) **Content features**

Semantic contents are less effective due to their variable dependency nature and are easily editable and controlled by the author. While semantic features are difficult to manipulate, they are therefore more useful than content features [32].

3 New and Current Feature Set

It is a mixture of different feature sets. It includes lexical, syntactic, content specific, and structural feature sets. This set includes the common word, misspelling, prefix, suffix, previous and next words of a certain word, *n*-grams categories, Lexical features syntactic features used with the part-of-speech tagger. Some studies used special, rare words, Binary Common Words Synonyms [28, 31]. In this category named entities, hapax legomena, gunning fog index, Flesch-Kincaid readability, syllable count, gunning fog index, smog index, delimiters, interjections, lexical density, hard words, unique words, adjective, preposition, noun, pronoun, conjunction, passive voice count are used.

4 Stylometry

Some stylists have explored a large variety of associated questions and research topics like [33] Are he/she the right author of this text/document? [34] Who is the genuine writer of the text/document among a group? [35] How many authors wrote a document? [36] In what mood was the document written? [36] Gender determination of

author through a document. [37] When was the document written? Has this document been translated? Stylometry has the power to disclose the original owner of a text document [2] used frequencies of FWs because function word represents the grammatical relationship among the words in a sentence and define syntactic relationship in the sentence. Since FWs are topic independent and consider as an efficient and effective scale for the AA. AA plays an essential role in the stylometry and computational analysis of writing techniques [25, 38–40]. Stylometric features generally involve the hidden clues of writing technique in a document, which appear unconsciously during document writing. Such features may be quantifiable and salient [4] and cannot be manipulated [19]. The precision of the result in AA depends on the methods, test and training data size, and the feature vector.

Stylometric techniques for analysis are generally divided into two parts supervised and unsupervised. Supervised techniques require author-class labels for classification, whereas unsupervised techniques do not require prior knowledge of author classes for classification. Supervised methods like support vector machines (SVMs) [41], neural networks [29, 41], decision trees [34], and linear discriminant analysis [20] for AA task. SVM is a strong classification technique. It provides powerful classification capabilities for online AA. SVM is similar to other supervised learning algorithms such as neural networks (NNs) and decision trees (DTs) in comparison [29, 34]. Among the unsupervised AA classification techniques, and Cluster Analysis and Principal Component Analysis (PCA) were used [31]. PCA captures essential features from a batch of feature vectors with low dimensionality which builds it more appealing for text attribution problems, it typically involves big size feature sets several authors used PCA in their work [19, 42], and has also proven effective for this, in unsupervised stylometric classification techniques, cluster analysis and principal component analysis (PCA) generally used [**Holmes 1992**]. PCA captures essential features from a batch of feature vectors with low dimensionality which makes it alluring for text analysis, it typically involves large feature sets in the stylometric analysis [37] (Table 1).

The work is summarized in Table 2.

Comparison of work done by the author in percentage: This comparison is the essence of (approximate 120) research papers.

5 Conclusion

In our study, we found that vocabulary, character features are more commonly used by researchers in their studies because they are more reliable. The functions word is just one attribute that is used extensively in AA, is a successful, most popular feature. The character *n*-gram features are robust, perform well for a variety of training, test data on different topics. Character *n*-gram can be good predictors. The combination of word, character level *n*-gram features gives good, accurate results due to their complementary nature. Lexical features can be applied to any corpus in any language.

Table 1 Features and methods used by researchers

Year	Authors	Features	Classifiers/methods
1887	Mendenhall	Sentence length, word length	Distance
1888	Mascolo	Function words, punctuation	Distance
1932	Zipf	Common word frequency, sentence length	Distance
1944	Yule	Vocabulary richness	K-means, distance
1952	Fucks	Word length	Distance
1956	Morton	Sentence length	Distance
1963	Brinegar	Word length	Distance
1964	Mosteller et al.	Function words	Naive Bayes
1965	Morton	Sentence length	
1987	Burrows	Function words	MVA, PCA
1992	Burrows	Function words	MVA, PCA
1993	Mathews et al.	Function words	Nearest Neighbor
1995	Holmes	Rewrite rule, POS, function words, character *n*-gram	
	Holmes et al.	Function words, vocabulary richness	MVA
	Kjell et al.	Character *n*-gram	NN, KNN
1996	Baayen et al.	Syntactic features	NN
	Tweedie et al.	Function words	NN
1998	Agramon et al.	Function words, POS, *n*-gram	ID3, Ripper
	Tweedie et al.	Function words, vocabulary richness	Distance, MVA, PCA
	Holmes	Word count begins with a vowel, bow	
1999	Binogo, Smith	Function words	MVA, PCA
2000	Stamatatos	Syntactic chunks	Distance
2001	Kukushina et al.	Character *n*-gram, POS	Markov Distance
	Chaski	Syntax, punctuation	Distance
	Holmes et al.	Function words	MVA, PCA
	Stamatatos	Syntactic features	Distance (LDA)
2002	Baayen et al.	Function words, syntax	MVA, PCA
	Khemelev et al.	Character *n*-gram	Markova Distance
2003	Binogo	Function words	MVA, PCA
	Clement et al.	Character *n*-gram	NB
	Hoover	Words, *n*-gram	MVA
	Keselj et al.	Character *n*-gram	MVA

(continued)

Table 1 (continued)

Year	Authors	Features	Classifiers/methods
	Khemelev, Teahan	Character *n*-gram	Markova Distance
	Agramon	Function words, POS *n*-gram	Winnow
	Koppel, Schler	Function words, POS *n*-gram, idiom	SVM, J 4.8
2004	Hoover	Words *n*-gram	MVA, PCA
	Peng et al.	Character *n*-gram, word *n*-gram	NB
	Abbasi, Chen	Characters, words	SVM, J 4.8
2005	Chaski	Character *n*-gram, word *n*-gram, POS *n*-gram	Distance (LDA)
	Juola, Baayen	Function words	Distance (cross-entropy)
	Zhao, Zobel	Function words,	NB, J 4.8, KNN
	Koppel	Function words, POS, *n*-gram, idiosyncrasies	SVM
2006	Koppel et al.	Function words, POS, punctuation	Balanced Window
	Zhao et al.	Function words, POS, *n*-gram	SVM, distance
	Zheng et al.	Characters, function words, syntax, vocabulary richness	NN, J 4.8, SVM
2007	Agramon et al.	Function words, syntax	SVM
	Burrows	Words	MVA, Zeta
	Hirst, Feiguina	Syntax	SVM
2008	Abbasi, Chen	Characters, function words, syntax, vocabulary richness	SVM, PCA
	Agramon et al.	Function words	Bayesian Regression
	Stamatatos	Character *n*-gram	SVM
2010	Kim Luyckx	Lexical, syntactic, character features	K-Nearest Neighbor
	Navot Akiva	Bow	SVM
2013	Sarwat Nizamani et al.	Lexical, syntactic, structural, content specific	Cluster-based classification
	Moshe Koppel et al.	Common words, character *n*-gram	Bayesian logistic regression, SVM
	Roy Schwartz et al.	Character *n*-gram features, word *n*-gram	SVM with a linear kernel
2014	Moshe Koppel et al.	Min–max matrix, bow, character-tetra grams	SVM
	Christian Mayor	BOW, character bi, tri-grams	Imposter, NB, Sparse

(continued)

Table 1 (continued)

Year	Authors	Features	Classifiers/methods
	Genard lynch	Character *n*-gram features, POS tags, type-token ratio, lexical richness, readability scores	SMO SVM, NB
	Imene Bensalem et al.	Character *n*-gram	Naive Bayes
	Jan Rygl	Lexical, character *n*-gram word *n*-gram	SVM
	Fatma Howedi et al.	Lexical, character *n*-gram word *n*-gram	Naive Bayes, Support Vector Machines
2015	Daniel Castro	Character, lexical, syntactic, semantic	SVM
	M. Sudheep Elayidom et al.	Lexical, syntactic features	Fuzzy, SVM
2016	Steven. Ding et al.	Lexical, syntactic, and character level	LDA, LSA, PVDM, PVDBOW
	Sebastian Ruder et al.	*N*-gram	CNN
2017	Prasha Shrestha et al.	Character *n*-gram	Convolutional Neural Networks
	David Kernot et al.	Richness, personal pronouns, referential activity power, sensory-based adjectives, words	PCS, LDA, and VSM
	Roy Schwartz et al.	Sentence length, character *n*-gram, word *n*-gram	Neural Network
	Jurgita kapociute-dzikiene et al.	Word or character *n*-gram	Naive Bayes Multinomial, SVM
	Stefan Evert, Thomas Proisl	Word frequency, syntactic features	Burrows's Delta measure
	Srinivasa Rao et al.	Word *n*-gram	*K*-means
	Jeevan Kumar et al.	Term weight measure, BOW	Naive Bayes
2018	Jagadeesh Patchala et al.	Part-of-speech, character *n*-gram, function words	SVM
	Paolo Rosso et al.	Character *n*-gram	CNN
	Oren Halvani et al.	Character *n*-gram	Nearest Neighbor, SVM
	Jagadeesh Patchala et al.	Syntactic features	Naive Bayes, SVM
2019	Biveeken Vijayakumar et al.	Word *n*-gram	Naïve Bayes, SVM
	Biveeken Vijayakumar et al.	Stylometric features, word, character *n*-gram	Naïve Bayes, SVM
	Nur Inda Jazilah	Fixed one-gram	KNN, SVM with Stylometric function
	Waheed Anwar et al.	*N*-gram	Latent Dirichlet allocation

(continued)

Table 1 (continued)

Year	Authors	Features	Classifiers/methods
	Raju Dara, Raghunadha Reddy	*N*-gram of character, word, POS	Rom forest classifier, Naïve Bayes Multinomial classifier
	Rea Bacciu et al.	Character *n*-gram, word *n*-gram, POS	SVM with RBF kernel
	Pian et al.	*N*-gram, POS, FW, SF	NB, SVM, Functional Tree
	Al-Falahi Ahmed et al.	Character *n*-gram, function words	NB, SVM, LDA
	Fatma Howedi et al.	Character *n*-gram, rare words frequency	K-Nearest Neighbor
	Michael Tschuggnall et al.	Distributed BOW	SVM
	Hans van Halteren	*N*-gram, POS	SVM
	Mostafa Rahgouy et al.	*N*-gram	Linear SVM
2020	Abeer H. El Bakly et al.	Word *n*-gram, term frequency bag of words	Ontology
	Mael Fabien et al.	Lexical *n*-grams, hapax legomena, content specific	Bidirectional Encoder Representations from Transformers
	Palacharla Ravikumar et al.	Word *n*-gram, POS	Naive, SVM
	Aisha Khatun et al.	Character *n*-gram	CNN
	A. Pian et al.	Stylometric features	Weka toolkit
2021	HainingWang et al.	Stylometric features	SVM, Neural Network
	Zhang et al.	Word *n*-gram, POS	Naive Bayes
	Apoorva, Sangeetha	Stylometry features	Deep Neural Network
	Aleks, Romanov et al.	Stylometry features	Deep Neural Network

Syntactic features have proven to be more reliable style markers than lexical features because they are not under the conscious control of the author, cannot be easily manipulated. Content attributes are based on the semantic content of a document, are less effective because they are variable dependent in nature, easily editable, under the conscious control of the author. Recent research has also introduced some new features, this set includes common words, misspellings, prefixes, suffixes, previous, next words of a particular word, *n*-gram categories. One type of feature set is not sufficient for authorship attribution, mixing or combining two or more feature types gives accurate results.

Table 2 Show the comparison of features used by the authors

Features	Used by %	Advantages, disadvantages
Lexical features, e.g., word frequencies, word *n*-gram, function words, hapax legomena, word/sentence length, etc.	Approx. 99%	Almost every researcher used this feature, without this feature AA seems impossible
Syntactic features, e.g., POS, rewrite rule	20–30%	This feature is impossible to modify because it is an unconscious feature, the habit of a writer so inevitable
Content-based features, e.g., comma, quotation marks, question words, etc.	10–15%	Easily catch the writing style of the author. Easily editable
The new, current feature set, e.g., common word, misspelling, prefix, suffix, previous, next words of a particular word, rare words gunning fog index, Flesch-Kincaid readability, syllable count, gunning fog index, smog index, delimiters, interjections, lexical density, hard words, unique words, adjective, preposition, noun, pronoun, conjunction, passive voice count, etc.	5–10%	This is very useful with lexical, syntactic features proved in recent work. But without the first two features, it cannot imagine alone. More dependent on other features. Not used by more researchers because these are the latest features

References

1. Dabagh RM (2007) Authorship attribution and statistical text analysis. Metodoloski zvezki 4(2):149–163
2. Mosteller F, Wallace D (1964) Inference and disputed authorship: the federalist. Series in behavioral science: quantitative methods edition. Addison-Wesley, Reading
3. Peng RD, Hengartner NW (2002) Quantitative analysis of literary styles. Am Stat 56(3). https://doi.org/10.1198/000313002100
4. Holmes G, Nevill-Manning CG (1995) Feature selection via the discovery of simple classification rules. In: Proceedings of the symposium on intelligent data analysis, Baden, Germany. Holte RC. Very simple classification rules perform well on most commonly used datasets. Mach Learn 63–91
5. Stamatatos E (2007) Author identification using imbalanced and limited training texts. In: Proceedings of the 18th international conference on database and expert systems applications. IEEE Computer Society, Regensburg, pp 237–41
6. Stamatatos E (2009) A survey of modern authorship attribution methods. J Am Soc Inform Sci Technol 60(3):538–556
7. Schwartz R, Sap M, Konstas I, Choi Z, Li Y, Smith Noah A (2017) The effect of different writing tasks on linguistic style: a case study of the ROC story cloze task. In: Proceedings of the 21st conference on computational natural language learning, CoNLL
8. Juola P (2016) Future trends in authorship attribution, IFIP international federation of information processing volume 242. Adv Digit Forensics 3:119–132
9. Koppel M, Akiva N, Dershowitz I, Dershowitz N (2011) Unsupervised decomposition of a document into authorial components. In: Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies, vol 1. Association for computational linguistics (2011), pp 1356–1364

10. Koppel M, Schler J, Argamon S (2010) Authorship attribution in the wild. Lang Resour Eval. (Advanced Access published). https://doi.org/10.1007/s10579-009-9111-2
11. Argamon S, Levitan S (2005) Measuring the usefulness of function words for authorship attribution
12. Burrows J (2007) All the way through testing for authorship in different frequency strata. Literary Linguist Comput 27–47
13. Ding SHH, Fung BCM (2016) Learning stylometric representations for authorship analysis. Inst Linguist Evid [ILE]
14. Houvardas J, Stamatatos E (2016) N-Gram Feature Selection for Authorship Identification. Working notes papers of the CLEF
15. Keselj V, Peng F, Cercone N, Thomas C (2003) N-gram-based author profiles for authorship. In: Proceedings of the 6th conference of the Pacific association for computational linguistics. Halifax, Canada: Pacific association for computational linguistics (2003), pp 255–64
16. Sapkota U, Bethard S (2015) Not all character N-grams are created equal: a study in authorship attribution. In: Human language technologies, the annual conference of the North American chapter of the ACL, Denver, Colorado, pp 93–102
17. Luyckx K, Daelemans W (2012) The effect of the author set size and data size in authorship attribution. J Literary Linguist Comput 26(1):35–55. ISSN: 0268-1145
18. Stamatatos E (2008) Author identification: using text sampling to handle the class imbalance problem. Inf Process Manage 44(2):790–799
19. Baayen H, Van Halteren H, Tweedie F (1996) Outside the cave of shadows: using syntactic annotation to enhance authorship attribution. Literary Linguist Comput 11:121–31
20. Baayen H, Van Halteren H, Neijt A, Tweedie F (2002) An experiment in authorship attribution
21. Kukushkina OV, Polikarpov AA, Khmelev DV (2001) Using literal and grammatical statistics for authorship attribution. Probl Inf Transm 37(2):172–183
22. Cranenburgh V (2012) Literary authorship attribution with phrase-structure fragments. In: Workshop on computational linguistics for literature, pp 59–63, Montréal, Canada
23. Koppel M, Argamon S, Shimoni A (2003) Automatically categorizing written texts by author gender. Literary Linguist Comput 401–12
24. Feng S, Banerjee R, Choi Y (2012) Characterizing stylistic elements in syntactic structure. In: Proceedings of the joint conference on empirical methods in natural language processing and computational natural language learning (2012) (Anthology: D12–1139; Volume)
25. Argamon S, Whitelaw C, Chase P (2007) Stylistic text classification using functional lexical features. J Am Soc Inf Sci Technol 802–22
26. Stamatatos E, Fakotakis N, Kokkinakis G (2000) Text genre detection using common word frequencies. In: Proceedings of the 18th international conference on computational linguistics, vol. 2. Association for Computational Linguistics, Saarbrucken, Germany, pp 808–14
27. Argamon S, Saric M, Stein SS (2003) Style mining of electronic messages for multiple authorship discrimination: first results. In: Proceedings of the association for computing machinery conference on knowledge discovery and data mining [ACM SIGKDD], pp 475–480
28. Stamatatos E (2009) A survey of modern authorship attribution methods. J Am Soc Inform Sci Technol 60:3
29. Stamatatos E (2017) Authorship attribution using text distortion. In: Proceedings of the 15th conference of the European chapter of the association for computational linguistics: volume 1
30. Argamon S, Whitelaw C, Chase P, Dhawle S (2005) Stylistic text classification using functional lexical features (2005)
31. Holmes (1994) Authorship attribution. Comput Humanit 28(2):87–106
32. Baayen H, Van Halteren H, Tweedie F (1996) Outside the cave of shadows: using syntactic annotation to enhance authorship attribution. Literary Linguist Comput 11:121–131
33. Akiva N, Koppel M (2012) Identifying distinct components of a multi-author document, pp 205–209
34. Altheneyan S (2014) Naıve Bayes classifiers for authorship attribution of Arabic texts. J King Saud Univ Comput Inf Sci 26:473–484

35. Abbasi A, Chen H (2008) Writeprints: a stylometric approach to identity-level identification and similarity detection in cyberspace. ACM Trans Inf Syst 26:1–29
36. Abbasi A, Chen H (2005) Applying authorship analysis to extremist group webforum messages, Published by the IEEE Computer Society
37. Abbasi A, Chen H (2006) Visualizing authorship for identification. In: Mehrotra S et al (eds). ISI 2006, LNCS 3975, p 60. Springer, Berlin
38. Halteren V (2007) Author verification by linguistic profiling: an exploration of the parameter space. Assoc Comput Mach Trans Speech Lang Process 4(1):1–17
39. Juola P (2006) Authorship attribution. Inf Retrieval 1(3):233–334
40. Koppel M (2008) Computational methods in authorship attribution. JASIST 60(1):9–26
41. Diederich J, Kindermann J, Leopold E, Paass G (2000) Authorship attribution with support vector machines. Appl Intell 109–23
42. Burrows J (2002) Delta': a measure of stylistic difference and a guide to likely authorship. Literary Linguist Comput 16:267–87

Is Data Science and Blockchain a Perfect Match?

P. Swathi, M. Venkatesan, and P. Prabhavathy

Abstract Blockchain technology can help with IoT, digital transaction records, and supply chain management. Enterprises over the world are focusing more on the innovative work of blockchain technology. Blockchain innovation permits the end-client to do exchanges with each other without the association of the middle man. Any benefits from cash to motion pictures can be moved, put away, executed, oversaw, and traded with no intermediaries. Data science, as we know it, is the science to extract valuable insights and information from both structured and unstructured data to solve real-world problems. With a lot of advantages and challenges, blockchain and data science can end up being an amazing blend to oversee information amount with quality productively. In this paper, we are investigating the new aspect of mixing data science and blockchain.

Keywords Blockchain · IoT · Digital transaction · Supplychain management · Information · Data science

1 Introduction

Blockchain and data science are among the top arising advancements tipped to alter a few enterprises, profoundly changing the manner in which organizations and associations are run. One may accept that these innovations are fundamentally unrelated, each producing remarkable ways and applied independently of each other.

Data science, as we probably are aware, is the science to take out significant knowledge and data from both structured and unstructured information to solve real-world problems. The development of data science can be credited to the transient

P. Swathi (✉) · M. Venkatesan
National Institute of Technology, Karnataka, Mangaluru, India
e-mail: swathip.187co005@nitk.edu.in

P. Prabhavathy
Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: pprabhavathy@vit.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_9

ascent of enormous information. Huge information manages humongous volumes of information that regularly cannot be overseen by ordinary data handling techniques. With the appearance of enormous information, associations can store huge data. Data science empowers organizations to settle on better choices and expectations by finding concealed information designs from crude information. It is tied in with getting information experiences from the historical patterns that uncover various information points, which may be obscured prior.

Blockchain is a disseminated record comprising of various hubs associated without a central authority. This record is spread over a scope of PCs over the globe and run by anybody with an Internet association. Consistent with its name, blockchain innovation spins around the idea of a progression of interconnected blocks that structure a chain.

Nowadays, researchers are utilizing blockchain innovation to guarantee genuineness and track the information at each point on the chain. One reason for its huge acceptance is its vast range of security. With blockchain's decentralized record, information is ensured at each progression through different digital signatures. Blockchain, simply like data science, is a step-by-step transformation in which a few enterprises work. While data science centers exploit information for legitimate organizations, blockchain guarantees the trust of information by keeping up a decentralized record.

1.1 Scope of the Paper

Several works have been done by researchers on various aspects of adoption of blockchain in some particular data science techniques. According to our survey, most of the papers dealt with specific application or field where blockchain can club. There are only few papers which talks about the incorporation of blockchain with data science. A survey led by Meng et al. [1] portrays how blockchain helps to meet intrusion detection. The authors relieve trust issues by building up a synergistic IDS utilizing smart contracts [2]. Salah et al. [3] and Casino et al. [4] led an audit on blockchain-based applications. Ucci et al. [5] investigated malware examination utilizing ML methods. Highlights of malware were altogether talked about, and a point-by-point scientific categorization has been proposed. Similarly, Conti et al. [6] studied privacy and security issues of bitcoin. They talk about different classes of assaults, such as double spending assaults, customer side security dangers, and mining pool assaults. To deal with the previously mentioned issue, Rahouti et al. [7] examine specific ML-based arrangements and examine the treatment of specific social issues, for example, human trafficking and drug deals through digital currency.

The remainder of the paper is as follows. In Sect. 2, we discuss some of the background knowledge, such as data science and blockchain. In Sect. 3, we explain an overview of integrating data science and blockchain. In Sect. 4, we highlight the impact of blockchain on data. In Sect. 5, we discuss the use cases of blockchain and data science. Finally, in Sect. 6, we conclude the paper.

2 Background

2.1 Data Science

Data science is a "concept to unite statistics, information analysis, and their related techniques" in order to "understand and study true wonders" using data. It is strategizing and hypotheses from a variety of disciplines, including computer science, mathematics, information science, statistics, and domain expertise. Szalay [8] envisioned data science as a "fourth worldview" of science (experimental, hypothetical, computational, and now information-driven) and attested that "everything about science is changing as an effect of data innovation." As the world entered the period of quite large information, the requirement for its stockpiling additionally developed. Previously, it was the major test and source of concern for undertaking businesses. The focus was mostly on developing a data storage system. When Hadoop and other architectures have successfully addressed the issue of capacity, the focus has shifted to the processing of this data. The key component here is data science. Data science can turn all of the images from Hollywood science fiction flicks become a reality. Artificial intelligence's future will be data science. As a result, it is critical to understand what data science is and how it may help the corporate sector become more valuable [9].

Predictive analysis is required whenever you need a model that can forecast the possible outcomes of a given event in the future. You may create a model that can do predictive analysis on your work in order to anticipate whether or not the future will be on time. If we want a model that can make its own decisions and change them with dynamic bounds, prescriptive analytics is the way to go. Giving guidance is essential to this relatively new industry. It both predicts and suggests a range of recommended actions and outcomes. Google's self-driving vehicle is the finest model for this car. Automobiles can collect data that can be used to prepare self-driving vehicles. It can do computations on this data in order to transfer knowledge to it. This gives the vehicle the ability to make decisions such as when to turn, which way to stop, and whether to back off or accelerate. If you do not have the parameters on which to base a forecast, you will have to look for patterns within the dataset to create meaningful predictions. Because you do not have any specified labels for grouping, this is just the model. Some well-known methods of grouping include classification and clustering.

2.2 Blockchain

Blockchain is a decentralized computational [10] and information sharing platform that allows different authoritative domains that do not trust each other to interact, coordinate, and participate in rational decision-making. It is an electronic, decentralized ledger that maintains track of all peer-to-peer transactions. It is a continuous list of transaction records stored in encrypted form, which are interlinked to each other

like a chain. Each block is uniquely connected with the previous block by a digital signature so that record cannot be altered or tampered without disturbing the records in the previous block of the chain. The unique feature of blockchain is that there is no need for a third-party authentication mechanism. When the whole peer-to-peer network agrees on a transaction, it is considered legitimate. Bitcoin, a cryptocurrency, is one of blockchain's applications. The hash tree, also known as a Merkle tree, was the most basic version of the blockchain. The hash tree data structure was developed by Ralph Merkle in 1979, and it was used to validate and handle data across computer systems. Validating data is crucial in a peer-to-peer network of computers in blockchain to ensure that nothing was altered or modified during the transfer. It also made it possible to share only authenticated data across the network. In a nutshell, it is utilized to keep and validate the integrity of data that is being exchanged. The Merkle tree was first used in 1991 to build a "secured chain of blocks"—a sequence of data entries that were all linked to each other. The most recent record in this chain would include the whole chain's history. Each block has its own encrypted data as well as some information from the preceding block. Satoshi Nakamoto is the pseudonymous person or individuals who invented bitcoin, wrote the bitcoin white paper, built and implemented bitcoin's first reference implementation, and envisioned the distributed blockchain in 2008. It would keep a secure record of data transfers, use a peer-to-peer network to time stamp and authenticate each transaction, and be self-managed without the need for a central authority. Bitcoin's foundation was built on this. As a result, the blockchain, and also the world of cryptocurrency, emerged.

Let us see the working of blockchain in terms of the Bitcoin transaction life cycle. Whenever Alice wants to send some coins to Bob, initially, Alice opens her Bitcoin wallet and provides the address of Bob and the amount to transfer. Then, he presses the send button. Once Bob presses the send button. Then, the wallet constructs the transaction a transaction like that, and this transaction is signed using Alice's private key. By applying the digital signature techniques wallet signs the transaction made by Alice and broadcast it over the network. All hubs in the system or the majority of the hubs in the system receive that particular transaction. After receiving the transaction, the nodes in the network will validate the transaction based on the existing blockchain. Once this transaction is validated, then it is propagated to some particular nodes called miners [11]. The miner collects all the transactions for a time duration, and they construct a new block and tries to connect it with the existing blockchain through some cryptographic hash computation. Miners take this computationally tricky hash problem and try to solve it. Whoever solves it can connect that block with the existing blockchain. They propagate the updated blockchain in the network.

The data models used in blockchains vary from key-value to document stores and are usually combined with "off-chain" data stores. Consequently, looking and recovering heterogeneous information in blockchain-based frameworks takes handmade and impromptu programming endeavors, not at all like the theoretical and explanatory question procedures in customary datasets. Considering the expanding interest for blockchain information examination at scale, seeing how to proficiently access, incorporate, and break down information in this heterogeneous climate is impor-

tant. The volume of information that blockchain networks store and oversee will just develop with time. However, numerous contemporary executions show low throughput, low versatility, and high inertness. Also, to balance the significant expense of building trust among executing parties through agreement, just as to debilitate torpid information, charges are charged in open blockchains for both putting away and controlling information. Appropriately assessing the on-chain/off-chain information structural decisions of a blockchain application can help settle a portion of these issues. The data put away in blockchains are lasting and open to the entire organization. This achieves a scope of information administration issues, for example, security and quality confirmation. While putting away the information in scrambled structure is suggested, it may be dependent upon brute-force decryption attacks later on or lead to unintended security spillages. Hence, it is basic to painstakingly audit these issues to help create sufficient structures for blockchain information administration to advance viable administration and appropriate utilization of blockchain innovation.

3 Integrating Blockchain and Data Science

Since we have top to bottom data about blockchain and data science, it is evident that these are two unique innovations with various objectives. At the same time, data science means to encourage information examination for significant bits of knowledge and better decision-making; blockchain centers around recording and approving information. Both of these advancements use algorithms to accomplish what they are proposed to do.

To sum up, data science empowers information to forecast, while blockchain guarantees information trustworthiness. Along these lines, on the off chance that we are contrasting them, it resembles contrasting one type with a totally different type. In any case, on the off chance that we use their pair, they can give valuable treasures.

Some factors that will empower the combination of blockchain and data science are data integrity, data security, prediction, data sharing, and data analysis. Data integrity alludes to the unwavering quality and trustworthiness of the information. It includes the maintenance and the confirmation of the exactness and consistency of information over its whole life cycle. Blockchain guarantees data integrity with its encryption and rigid verification techniques. Further, it gives truly necessary straightforwardness through transaction discernibility. Blockchain might even be the solution to achieving the highest levels of data integrity. Blockchains are designed to be immune to data manipulation. Blockchain records are immutable, which means that once information has been expanded or exchanged, it cannot be changed or destroyed. It is there, and it will be there. Furthermore, blockchains are an information structure as well as a timekeeping instrument for the data. So the proof of the historical backdrop of information is effectively reportable and updated. Associations confronting a review, regulatory prerequisites, or legal difficulties can utilize blockchain innovation to improve data integrity.

The "Decentralization" of blockchain makes it hard for programmers to assault the data as it will require trading off all the hubs, which is essentially unimaginable [12]. Further, blockchain naturally removes any hub that carries on dubiously, making the framework secure. Blockchain guarantees that your information is cryptographically encrypted, which implies that change in information is a troublesome errand. You can likewise spare a cryptographic mark of an archive or record on a blockchain. This would give clients an approach to guarantee a document is unaltered without expecting to spare the whole record on the blockchain. Due to its decentralized nature, you can generally cross-check document signatures over all the records and confirm that they have not been changed. At the point when you take a look at a record, you can ensure that it is the very same file that existed before. In the event that somebody changes a record, at that point, the signature is delivered invalid. Blockchain offers dependable, autonomous information check which is unquestionable.

Data science's ability can be used to dissect blockchain information to infer significant information experiences and hidden patterns. By utilizing a blockchain organization to store the data from studies, researchers can forestall using previously utilized information or try not to rehash information examination that is now led before. The innovation can help send the data safely without the need to duplicate data cleansing. While blockchain offers constant exchanges, data science gives top to bottom data analysis. These two technologies can be joined to convey continuous data analysis that can alter numerous ventures and smooth out business measures.

With the combination of advantages and challenges, blockchain and information science can end up being an amazing blend to oversee information amount with quality proficiently. More advancements and development of blockchain innovation will encourage the investigation of more use cases, including data science. Then again, data science can help blockchain with its low stockpiling cost. It will be intriguing to perceive how these advances develop to address current difficulties and grandstand their capability to change the information the executives and utilization.

4 Impact of Blockchain on Data

Data science, much the same as any innovative progression, has its own difficulties and constraints which, when tended to, will release its full abilities. Some significant difficulties to data science incorporate inaccessible, privacy, and security issues, and dirty data. The control of dirty (or incorrect data) is one portion that blockchain innovation can decidedly affect the data science field in no little measure. The consideration of messy information like a copy or inaccurate information was distinguished as the greatest test to data science. Through decentralized agreement calculation and cryptography, blockchain approves information making it practically difficult to be controlled because of the immense measure of computing power that will be required. Again through its decentralized framework, blockchain innovation guarantees the security and protection of information. Most information is stored in servers that are

frequently the objective of digital aggressors; the few reports of hacks and security breaks go to show the degree of the danger. Blockchain, then again, reestablishes the control of information to the people producing the information, making it a tough assignment for cybercriminals to get to and control information for an enormous scope. Blockchain has brought a totally different method of overseeing and working with information. Neither in the central point of view where all information should be united but a decentralized way where information might be dissected right off the edges of individual gadgets. Blockchain incorporates other cutting-edge innovations, similar to artificial intelligence (AI), cloud computing, and the Internet of Things (IoT). Besides, validated data produced by means of blockchain innovation comes organized and complete in addition to the reality that it is permanent. Another significant area where blockchain-created data turns into a lift for enormous information is data integrity since blockchain discovers the origin of data; however, it is connected chains.

5 Use Cases

The use of (i) DLT [10], (ii) cryptographical verification, and (iii) the ability to have smart contract reasoning embedded into it are instances of specific highlights of blockchains. This means that blockchains can allow many untrustworthy clients to execute directly on a record (smart contracts self-execute trades) without the need for a trusted intermediary. Transparency, decentralization, automation, and immutability are some of the innovative characteristics of blockchain technology. These divisions may be used for a variety of businesses, resulting in several application cases.

5.1 Banking

According to Danger et al. [13], banks must lend, but only to borrowers who are at grave risk. This motivates banks to collect detailed, individually identifiable information from everyone who applies for a loan, such as annual salary, date of birth, voter ID, Aadhar card, or passport information. Finally, banks use this information to determine a candidate's creditworthiness. Specific data may be required to be shared with experts under certain circumstances, such as to prevent unlawful tax evasion. However, with so much personal information on hand, each bank becomes a tempting target for hackers. Instead of disclosing any confidential information, loan applicants can construct ZKPs stating that their past year income charges are above a certain threshold that they have a valid government identification number, and that their financial evaluation exceeded a certain threshold within the preceding week. Distributed ledger identification creates a global source of truth, which benefits a large number of people. Consents can be granted by candidates, and everyone can agree on how and when it was given. Banks can make adjustments to rules and keep

a record of their actions. As a result, the market will be more efficient: banks will be able to provide loans with more precision, and applicants will be able to protect their personal information more effectively.

5.2 Healthcare

Enrollment and retention are the most challenging aspects of clinical trials, and despite several attempts over the years, progress remains mostly hidden. For these use cases, blockchain offers clinical preliminary data exchange and the ability for study individuals to experience esteem disclosure (counting individual well-being data adaptation). Patient data management is one of the most common blockchain use cases in human services. Health organizations will separate therapeutic documents [12], making it difficult to pick a patient's remedial history without contacting their previous care provider. This approach can take a long time to complete and is prone to errors owing to human error. Identity management of individuals (e.g., doctor, patient, member, and supplier), unique device IDs for restorative gadgets in the medical production network, or authoritative members or validators in a system is the essential aspect of blockchain use case in medicinal services. The most common use case for DLT across organizations is item supply, from inception to end-of-life. Blockchain is being used in social insurance, medicine, clinical supplies, blood products, and therapeutic devices for activities, consistency, and gauging among pharmaceutical producers, suppliers, pharmacy retailers, blood donation centers, and payers.

5.3 Agriculture

Before it reaches the buyer, the typical agriculture store network involves complicated, interrelated operations between many parties, such as the farmer, distributor, processing agencies, examination, and insurance agencies, planning and transit organizations, banks, and vendors. There are a few challenges in this lengthy operation [14]. Distributors believe it is difficult to track provenance in order to determine the origin and type of imported goods. The custody or custodial details becomes difficult to follow as objects travel between spouses. Global traders are wary of centralized entities, especially private organizations that certify crops. The data stream among stakeholders is causing a potential delay in fundamental leadership downstream. These problems can be solved with blockchain technology. Ranchers may suffer significant agricultural losses as a result of natural disasters such as floods, torrential rain, wind, earthquake, and landslide. When the crops are ruined, they must appeal to the government for compensation using a complicated procedure, and the approval is subject to a number of inspections. This architecture avoids complications by com-

bining blockchain technology. Customers may also use blockchain data to pay tips directly to the farmers who have provided them with the best service.

5.4 Supply Chain Management

The industrial network that transports fish from the ocean to the table is extremely complex and opaque. It has a large number of members from diverse businesses and administrative controls that span national borders. As a result, this business structure is an appropriate gateway for blockchain advancements [15]. Oceana, a non-governmental organization dedicated to protecting the oceans, theorized that a common platform for fish recognition would increase labeling precision and reduce private fishing: "Despite significant challenges, fish discernibility is well within reach." We can make ground against offshore trawling simply by tracking where our fish come from at each stage of the production chain. When a fish is captured, sensors are attached to it to record information such as the location from where it was caught, temperature, and moisture content. This information, as well as other events in the treatment of the fish, is recorded in the blockchain: stockpile temperature runs, ownership transfers, transportation organization, and so on. The record may also be used to investigate both administrative and logical aspects of fish harvesting and use.

5.5 Others

Blockchain has a wide range of applications in the commercial world, including aviation, telecommunications, and IoT [16]. Several features of blockchain must be embraced by many companies. Among these, decentralization is critical. A concentrated server is used in business to store all of an organization's data. Anyone who wants access to the data consults the database about the various levels of data. Because everything is dependent on the server, it is important to keep it running at all times. There will be congestion at some point. A single server failure will destabilize the entire system [17]. Unchanging nature is another aspect of a blockchain that companies must accept. There will be a full transformation of reality, which anybody may check to see whether the change in information causes a change in the hash value, which can detect the change. As a result, no one can change the information about the company. To put it another way, we may say that everyone is looking at the data. Therefore, no one will try to change it. The blockchain's third characteristic is its cost-effectiveness—the benefits of blockchain cause it to adapt in the corporate world [18]. The elliptic curve digital signature computation is used to carefully mark blockchain trades. The distribution of exchanges between hubs effectively demonstrates their origins. The validation process is computationally challenging and is the critical bottleneck. When an association has been created via separation in concentrated databases, there is no need to examine every sale that

arrives across it separately. The transaction's independent processing takes place in a separate location. Depending on the consensus method, this might include objective forward and backward communication [19], as well as the oversight of forks and the associated rollback.

6 Conclusion

Blockchain in its early stages may not show up so because of the promotion the innovation has gotten in a brief period. One would expect that as the innovation develops and there are more advancements around it, more solid use cases will be distinguished and investigated as data science is one area that will profit from this and concluded the paper at last.

References

1. Meng W, Tischhauser EW, Wang Q, Wang Y, Han J (2018) When intrusion detection meets blockchain technology: a review. IEEE Access 6:10179–10188
2. Xiong W, Xiong L (2019) Smart contract based data trading mode using blockchain and machine learning. IEEE Access 7:102331–102344
3. Salah K, Rehman M, Nizamuddin N, Al-Fuqaha A (2019) Blockchain for AI: review and open research challenges. IEEE Access 7:10127–10149
4. Casino F, Dasaklis T, Patsakis C (2018) A systematic literature review of blockchain-based applications: current status, classification and open issues. Telemat Inform 36:55–81
5. Ucci D, Aniello L, Baldoni R (2019) Survey of machine learning techniques for malware analysis. Comput Secur 81:123–147
6. Conti M, Kumar ES, Lal C, Ruj S (2018) A survey on security and privacy issues of bitcoin. IEEE Commun Surv Tutor 20(4):3416–3452, 4th Quarter
7. Rahouti M, Xiong K, Ghani N (2018) Bitcoin concepts, threats, and machine-learning security solutions. IEEE Access 6:67189–67205
8. Szalay A, Gray J (2006) 2020 computing: science in an exponential world. Nature 440:413–414. https://doi.org/10.1038/440413a
9. Memon Q, Ahmed M, Ali S, Memon AR, Shah W (2016) Self-driving and driver relaxing vehicle. In: 2016 2nd international conference on robotics and artificial intelligence (ICRAI), Rawalpindi, pp 170–174. https://doi.org/10.1109/ICRAI.2016.7791248
10. Nakamoto S (2009) Bitcoin: a peer-to-peer electronic cash system. Available from: http://www.bitcoin.org/bitcoin.pdf
11. Swathi P, Modi C, Patel D (2019) Preventing Sybil attack in blockchain using distributed behavior monitoring of miners. In: 2019 10th international conference on computing, communication and networking technologies (ICCCNT), Kanpur. IEEE, pp 1–6
12. Kuo T-T, Kim H-E, Ohno-Machado L (2017) Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 24(6):1211–1220. https://doi.org/10.1093/jamia/ocx068
13. Nguyen QK (2016) Blockchain—a financial technology for future sustainable development. In: Proceedings—3rd international conference on green technology and sustainable development, GTSD 2016, pp 51–54
14. Kamilaris A (2018) The rise of the blockchain technology in agriculture and food supply chain. ResearchGate

15. Apte S, Petrovsky N (2016) Will blockchain technology revolutionize excipient supply chain management? J Excip Food Chem 7(3):76–78
16. Alam T (2019) Blockchain and its role in the Internet of Things (IoT). Int J Sci Res Comput Sci Eng Inf Technol 151–157
17. Wang H, Zheng Z, Xie S et al (2018) Blockchain challenges and opportunities: a survey. Int J Web Grid Serv 14(4):352
18. Kshetri N, Voas J (2018) Blockchain in developing countries. IT Prof 20(2):11–14
19. Ben Hamida E, Leo Brousmiche K, Levard H et al (2017) Blockchain for enterprise: overview, opportunities and challenges. Blockchain for smart transactions—IRT SystemX. View project Kei-Le'o Brousmiche. Available from: https://www.researchgate.net/publication/322078519

Review of Road Pothole Detection Using Machine Learning Techniques

Ankit Shrivastava, Devesh Kumar Srivastava, and Aditi Shukla

Abstract The paper is about detecting the number of potholes on the roads using digital image processing and machine learning. There is a dire need for improving the road conditions and quality of roads. We felt that there is a need for such a software because the road conditions in many areas in the country are not that good, and the reason for many road accidents in today's generation have been the bad conditions of roads. This software will help us know the presence and the quantity of the number of potholes on the roads. This data can be sent to the road repairing/maintenance teams. They will be able to improve/repair the roads knowing the exact number of potholes present on specific roads/localities. The motivation to do this project came from the greater social cause of improving the roads so that every regular day of a person is safer than today. The procedure that we used was Image preprocessing and then further on we used machine learning to get our desired results. The paper kept us dedicated to keep learning and to use the knowledge for good use. Using digital image processing we will program the software which reads the image and pre-processes it. Using machine learning we will train the software to recognize the potholes on roads. In this paper, we discussed to identify the potholes on the roads with the use of a video and a dataset of the images. The video helped us to create frames; each frame checked for the presence of a pothole. Once the pothole gets detected, we counted the number of potholes in all the frames created by our algorithm. Our aim is to ease the process to remove the potholes on the roads that will in turn help reduce the large number of fatalities and accidents.

Keywords Digital image processing · Machine learning · Dataset · Pothole · YOLO

A. Shrivastava (✉) · D. K. Srivastava · A. Shukla
SCIT, Manipal University Jaipur, Jaipur, Rajasthan, India
e-mail: ankitshrivastavaieee@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_10

1 Introduction

A ground surface gap, typically an asphalt pavement where damaged sections of the pavement have been eroded by traffic is essentially a Pothole. Water in the underneath soil texture and vehicles traveling over the affected area are typically the result. The underlying soil is first eroded by water, then the traffic fatigues and fractures the poorly supported asphalt surface in the affected area. To create a hole in the ground, continued road activity expels both asphalt and the core soil material. The majorly rural population lives in India. The vast majority of rural areas in India lack access to all-weather roads and hence have challenging monsoon times. The Rs 100,000 crore has been allocated for constructing and maintaining roads in villages by the government in its 13th five-year plan. At a macro level, the circumstances are not substantially modified. India is currently spending approximately 20,000–30,000 crore annually on the roads project.

The challenges of overload and inadequate maintenance are one of the key factors behind this massive road maintenance spending. The truth is that nearly 70% of the money that is spent on repairing roads is directly paid out by workers.

The magnanimity of the expenditure incurred to repair roads is alarming. There is a need of help to reduce the charges incurred on an exponential level. The aim of our research is to help solve this problem. The time taking manual work of searching and finding road faults will get replaced by fast and accurate machine work. Machine/Software will also be able to multitask and find designated areas to be repaired simultaneously. This will speed up the process and in turn help us in avoiding heavy charges for labor work. As presentation we follow the strategies to stop the formation of potholes (Fig. 1) are as: 1. Pavement survey for risk factors; 2. Offering appropriate systems for drainage; 3. Precautionary repair activities; 4. Control of utility cuts. There is a dire need for improving the road conditions and quality of roads. We felt that there is a need for such a software because the road conditions in many areas in the country are not that good, and the reason for many road accidents have been the bad conditions of roads. This software will help us know the presence and the quantity of the number of potholes on the roads. By this

Fig. 1 Potholes on Indian roads

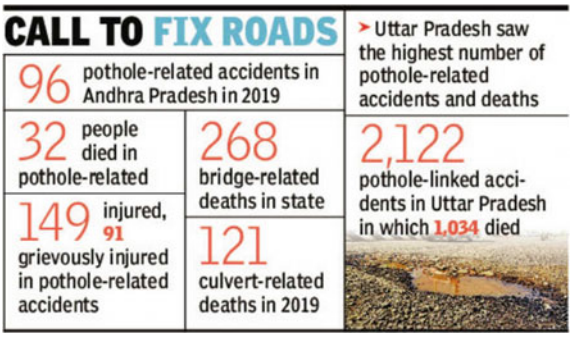


Fig. 2 Statistics of road accidents (internet resource)

data we can send the road repairing teams to specified locations to improve/repair the roads at those particular locations.

Referring to the above image one can see that the number of accidents (Fig. 2) due to the road conditions, road engineering/design are high. Our aim is to ease the process to remove such faults on the roads that will help us reduce this large number of fatalities. This will overall benefit the common man and the government overall. We will also provide solutions to the queries about potholes. This project will provide one stop solution to the road maintenance department for their problems regarding raw materials and the number of potholes. The aim of this paper is to detect the number of potholes on the roads using digital image processing and machine learning which brings mainly for the benefit of all those laborers, contractors, middlemen, maintenance department, the government and the common man who need this kind of information to curb the problem of potholes.

2 Literature Review

Karmel et al. [1] discussed automated methods for the identification of road losses. The classification algorithm they use for road management is highly efficient. Writers also used classification machine learning algorithms and convex hulls for image treatment. This technique is used for the implementation of application weighting. A pothole is the type of road destruction caused by collisions, water and traffic. Many varieties of textures are seen on the road from many factories such as different areas of the map. Therefore, damage to the road is very difficult to define. Each picture on the road has exclusive characteristics such as the road form and the street width making defect identification tricky to apply. Various image processing methods and lots of features were used to solve this issue and implement machine learning techniques.

For grouping, various methods were used to identify road fractures such as methods of the experimental setup for application, and methods of machine learning. This paper is used as a guide.

Song et al. [2] described a cost-effective approach to detecting bogs and humps on highways and provided drivers with early notices to deter collisions and vehicle injury. For detection of potholes and bumps, ultrasound sensors are used. A low-cost model has been proposed for the analysis of 3-D floor pictures which uses a low-cost Kinect sensor that provides direct depth measurements to minimize device costs. They suggested an support vector machine (SVM) pothole detection system. This approach differentiates potholes from other faults, for example, cracks. In partial differential equations, the images can be segmented. The system trains the SVM with a series of pavement pictures to spot potholes. Some special equipment and instruments are required to detect the state of roads in the previous research which causes additional costs and disadvantages. In this case, they use a handheld device to accurately locate road manholes.

Ryu et al. [3] created a process that can locate and classify potholes automatically. Established methods for pothole identification is categorized into vibrating methods, 3-D (3-D) reconstruction and vision methods. To detect potholes, the vibration dependent technique uses accelerometers. A vibration-based framework provides the advantages of providing minimal storage systems that can be used in real-time. The vibration-based approach often results incorrectly. For example, the hinges and joints on the road can be seen as potholes while accelerometers cannot identify the potholes in the middle of a lane, as they are not struck by any wheel. 3-D laser scanner techniques are available in real-time for the identification of potholes. The laser scanning systems are therefore very complex. These studies are currently focused on testing. Stereo display methods require a high measurement commitment to replicate floor surfaces by comparing feature points between two views. Therefore in a real-time environment, it is impossible to use them. The pavement depth was also collected using a low-cost Kinect camera. They had worked with hardware, and we may use their idea/method in future but it does not have any resemblance to our project.

Prathiba et al. [4] described a new structure for the segmentation of road photos that can distinguish the following topics in a hierarchical way: roads and slopes with or without the fall, mountains, road signs, vehicles, buildings and foliage. Experiments illustrate the solution to different road photos in this article is adequate. The more complicated roads then the organized roads are unstructured. To segment thin structures (cracks) in road surface photos, a multidisciplinary approach based on Markov's random field is proposed. A Gaussian 1-D smoothing philter strengthens the cracks and then processes them with a 2-D paired philter. For scientific purposes, a total of 64 photos of road pavement illustrating multiple types of cracks create a qualitative assessment. A transformation of non-sampled contour-let is used to diagnose defects, in which a small range of experimental findings is introduced. A comprehensive methodology for the automated identification and characterization of pavement defects is proposed by using line-based grayscale images which scan laser-lit cameras using the high-speed image acquisition system during road surveys.

The crack detection uses conditional texture anisotropy for each image and a multilayer neural perceptron network with two hidden layers for defect characterization. The findings presented are positive, but the examination of several splits on the

same picture does not help the distinction. Automated identification of pavement defects will recognize defects of crack type. A neural network MLPNN is used to identify the identification. To divide the images into four groups, neural networks are used: defect-free, crack, joined and bridged. The technique for path detection is built by integrating the invariant lighting function space and the classifier dependent on probability [5, 6].

3 Methodology

At first using digital image processing we will create a software which reads an image. Using machine learning we will train the software to recognize such potholes on roads. There will be a software tool where the user will feed a video consisting of a road with potholes. The software will process the video to give the desired results about the number of potholes present on the road. The pothole detection method that we have proposed is the image recognition method. This method will be carried out in three parts. The concepts that we are using are Image processing and machine learning.

Image processing.

The process to apply image preprocessing on your dataset is given below. The steps to be taken are:

- Read image
- Resize image
- Remove noise (De-Noise)
- Segmentation
- Morphology (Smoothing Edges).

At first our primary goal is to convert video into frames. These frames will consist of images of roads with and without potholes. The software needs to recognize the features of the image/frame. The main reason behind this is that the software needs to recognize a pothole irrespective of its shape and size. Using digital image processing we will write a program which reads and images. With the help of OpenCV we will be performing image preprocessing. Using machine learning we will train the software to recognize such potholes on roads. The final project will be a software where the user will feed a video consisting of a road with potholes to the software. The software will process the video to give the desired results about the number of potholes present on the road.

The two major parts of this domain is:

Video to Frame conversion.

Image Enhancement.

Image Preprocessing

This is an important step in image processing, it can reduce our dataset and therefore leads to an efficient analysis. The images in the dataset are RGB, but the pothole has a feature that it does not have the color of RGB. We will be using the OpenCV library to perform certain steps on the frames that were generated from the video to frame conversion. Converting RGB image to grayscale converts the 3-D matrix to 2-D and can reduce the dataset to $1/3^{rd}$ of its original size.

The frames/images are loaded into an array. These images are then processed to further steps consisting of image resizing, grayscale, de-noising and thresholding. Once we get the pre-processed image, we perform blurring operation and edge detection on the same to extract the boundary of the pothole. The various other features are also extracted that will be further used in the future. The region of interest is captured. Once the image is converted into a single channel using grayscale processing we will perform objection using yolo algorithm.

Object Detection-Yolo treats an image as a regression problem, instead of evaluating it through a complex pipeline. It divides an image into S X S grid and each grid cells predicts bounding boxes and confidence score(c) for that bounding box. If the center of an object is in the grid, then that grid cell will be responsible for the prediction. If, suppose there is no object in the box then the score will be 0. During training time, the algorithm can predict many bounding boxes but only one box predictor will be responsible for the prediction. We select that bounding box by Intersection over Union (IOU).

Each bounding box contains 5 prediction:(x, y, w, h, c), x is the distance between the center of the image the extremities of the box along the horizontal axis, its value ranges from 0 to 1, y is the distance between the center if the object and extremities along y axis, its value also ranges from 0 to 1, w and h are the predicted width height of the object and their value can be more than 1 and the 5th component c is the confidence score which is a measure of how sure a model is that the box contains the object. And how accurate its prediction is. C represents intersection over union between the box and the actual truth. Besides this, each grid also predicts conditional class probabilities, we only predict one set of class probability per cell.

Image scaling—The image size in digital photography refers to a digital image scaling. In image technology, optical object magnification is called upscaling or improving resolution. The primitives which shape the image can be scaled using geometric transformations without losing the quality of the image if a vector graphic image is scaled. Where a raster picture is scaled, a new picture must be created with a greater or lesser number of pixels. As the pixel number reduces (scaled down), it typically leads to a noticeable loss of consistency.

Image Threshold—A simplistic but reliable method of partitioning an image into the field is called Image Thresholding. This technique of analyzing images is a method of fragmentation that segregates entities by transforming gray images into binary images. Image thresholds are most effective in high contrast images.

Image grayscale—A grayscale image is used in digital imagery, computer-generated graphics, and colorimetry, with a single object reflecting only a quantity of light. It just contains details about strength. Grayscale pictures are a type of monochrome black and white or gray, consisting mostly of dark colors. The distinction is between black at weakest intensity and white at a higher level (Figs. 3 and 4).

Machine learning is being used to detect the potholes in all the frames. At first, we trained the machine using the dataset we have. Then we test it on the acquired frames

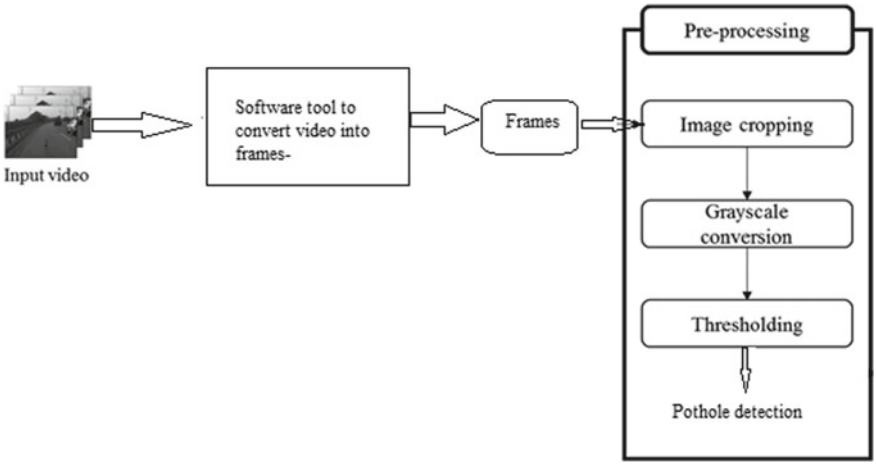


Fig. 3 Steps used in project

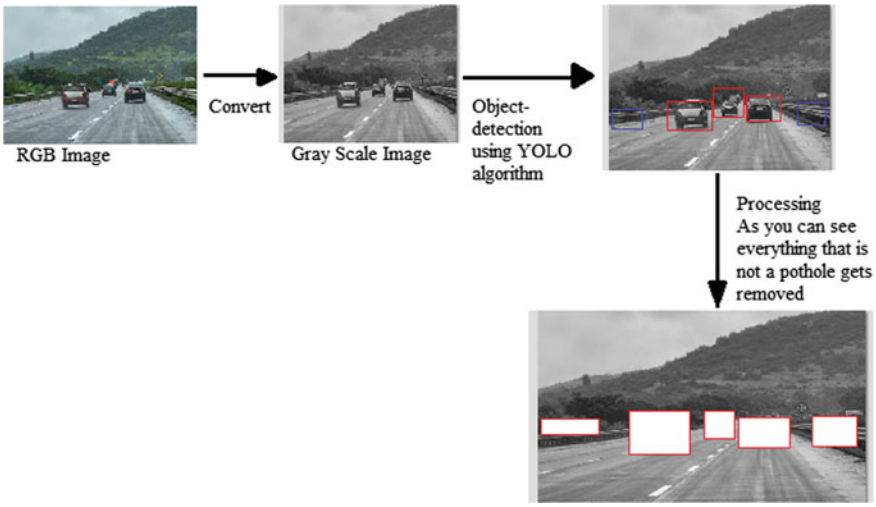


Fig. 4 No pothole detection after applying yolo algorithm

for the same. A support vector machine (SVM) is a discriminative classifier formally defined by a separating hyperplane. The algorithm outputs an optimal hyperplane which categorizes new examples. In 2-D space this hyperplane is a line dividing a plane in two parts where in each class lay in either side.

4 Implementation

It contains the screenshots of the code performed and the desired outputs (results) received. The implementation given below is in a sequence.

Algorithm for image extraction from videos

1. Step 1: Using Video Capture function of Open CV and passing the location of our video file. The video is converted into frames
2. Counting number of frames by get function store it in FRAMERATE
3. WHILE until the video is completed
    - 3.1 Capture the video frame by frame by read function and store it to RET and FRAME
    - 3.2 If there is no frame, i.e., if RET is false the loop gets terminated.
    - 3.3 To check if COUNT is the multiple of our FRAME RATE, then it is the frame we must save.
    - 3.4 To save the frame we will use in write function
    - 3.5 End WHILE
4. The images are detected from frames.
5. The images are pre-processed before execution (Fig. 5a–f)

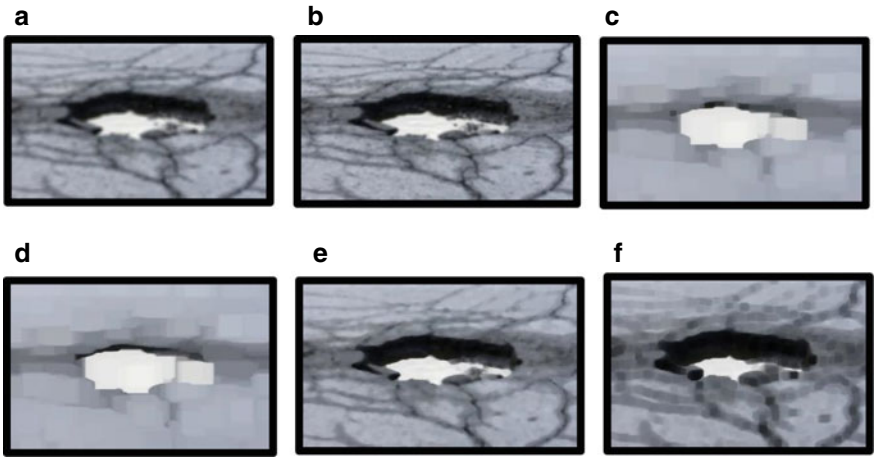


Fig. 5 **a** Blur. **b** Gaussian blur. **c** Median blur. **d** Erosion. **e** Erosion followed dilation. **f**. Dilation

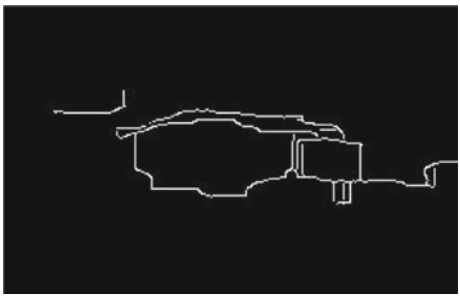


Fig. 6 Pothole detected

6. Apply the machine learning algorithm to detect the pothole from images (Fig. 6)

5 Conclusion

Potholes are of primary concern in road transportation. With this ever-increasing population and stagnant growth, this issue seems to be quite common in every city, this work is an attempt to contribute to a better infrastructure so that the problem can be resolved with minimum man power. We have implemented many algorithms and tested with many filters to reach the result. The work is a substantial progress in the field but there is still much scope of improvement.

We can build a system that can help us to have a snippet of the pothole, like in the application, we have images extracted from the video, similar techniques can be used for object detection and extraction. The conversion of the images retrieved from videos are converted into grayscale to reduce the computation time and speed up the object detection process. For the next step that is object detection, we used yolo algorithm to increase the accuracy of the pothole detection. Such potholes can be informed to PWD to repair these potholes so that the accident may reduce at high scale further we have used SVM to train the given model for the dataset and then test it on the acquired frames for the same. The algorithm outputs an optimal hyperplane which categorizes new examples. The paper proposed the challenge of predicting the actual size of the pothole, so that the material required for maintenance will be known as soon as the pothole is detected. There is a need for method that can maintain and manage roads continuously by detecting the pothole and also detecting the size and the equipment/material required for it.

References

1. Karmel A, Adhithiyan M, Senthil Kumar P (2018) Machine learning based approach for pothole detection. Int J Civil Eng Technol 18:882–888. (Author F, Author S (2016) Title of a proceedings paper. In: Editor F, Editor S (eds) Conference 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016))

2. Song H, Baek K, Byun Y (2018) Pothole detection using machine learning. Adv Sci Technol Lett 151–155. (Dept. of Computer Engineering, Jeju National University, Korea, Jan 2018)
3. Ryu SK, Kim T, Kim YR (2015) Image-based pothole detection system for ITS service and road management system. Math Probl Eng 2015. Article ID 968361:10p
4. Prathiba T, Thamaraiselvi M, Mohanasundari M, Veerelakshmi R (2015) Pothole detection in road using image processing. BEST Int J Manage Inf Technol Eng 3(4):13–20
5. Wu C, Wang Z, Hu S, Lepine J, Na X, Ainalis D, Stettler M (2020) An automated machine-learning approach for road pothole detection using smartphone sensor data. Published 2020 Sep 28 (Online) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7583950/. https://doi.org/10.3390/s20195564
6. Ochoa-Ruiz G, Angulo-Murillo AA, Ochoa-Zezzatti A, Aguilar-Lobo LM, Vega-Fernández JA, Natraj S (2020) An asphalt damage dataset and detection system based on retinanet for road conditions assessment. Appl Sci 10:3974. https://doi.org/10.3390/app10113974

Performance Improvement of the Multi-agent System of Knowledge Representation and Processing

Evgeniy I. Zaytsev, Elena V. Nurmatova, Rustam F. Khalabiya, Irina V. Stepanova, and Lyudmila V. Bunina

Abstract The article deals with the organization and development of the Multi-Agent System of Knowledge Representation and Processing (MASKRP). The diagrams of states, transitions and interaction of software agents, the types of queries to the knowledge base and the methods of knowledge processing are described. Ways to improve the performance of the Multi-Agent Solver and the construction of an optimal logical structure of a Distributed Knowledge Base (DKB) are proposed. In the algorithm for the synthesis of the optimal logical structure of the DKB, it is proposed to distinguish two interrelated stages. At the first stage, two problems are solved: the distribution of DKB clusters in the nodes of the computing system and the optimal distribution of data groups of each node by the types of logical records. At the second stage the problem of DKB localization on the nodes of the computer network is solved. As a result of optimization, the distributed knowledge base is decomposed into a number of clusters that have minimal information connectivity with each other. The results demonstrate the efficiency of the presented approach to development of the MASKRP.

Keywords Distributed systems · Multi-agent systems · Intelligent software agents · Data warehouse · Optimal logical data structure · Knowledge representation

1 Introduction

Multi-Agent System of Knowledge Representation and Processing belongs to a class of Distributed Artificial Intelligence Systems, in which a network of software agents is used to solve complex problems. The MASKRP uses two types of software agents: reactive and cognitive [1, 2]. The software agents of the MASKRP are able to take into account the so-called "NON-knowledge factors": inaccuracy, incompleteness, uncertainty, fuzziness of knowledge [3]. Interacting with each other through message

E. I. Zaytsev (✉) · E. V. Nurmatova · R. F. Khalabiya · I. V. Stepanova · L. V. Bunina
MIREA—Russian Technological University, Moscow, Russia
e-mail: zajcev@mirea.ru

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_11

passing, software agents find a solution to a complex poorly formalized problem. The development of special system software modules that implement inter-process communication mechanisms taking into account the specifics of the MASKRP allows to improve the system performance.

Traditional OS hide information about computing resources behind universal high-level abstractions, which makes it impossible for the MASKRP developer to implement problem-dependent optimization. To improve the system performance, it is necessary to be able to access hardware resources and implement specific mechanisms to control computing processes [4–11]. The efficient implementation of the MASKRP is possible on the basis of an exo-kernel Operating System and event-driven software modules that can directly interact with the low-level interface for accessing computing resources. Agent-oriented technology integrated with specialized system modules for the planning and control of computing resources becomes the main one for creating highly efficient knowledge processing systems [10–12]. Using the services of an exo-kernel Operating System, the developer of a Multi-Agent System has the ability to choose or implement their own System Libraries (LibOS). For example, the specialized Virtual Memory Management module and the Inter-Process Communication defined in LibOS can run significantly faster than universal software modules that do the same work in a monolithic or micro-kernel OS.

At the present stage, the idea of an exo-kernel architecture is implemented in the form of system virtualization using Virtual Machine Monitors (hypervisors). In this case, instead of a minimal exo-kernel that provides untrusted servers with a low-level interface for accessing computing resources, the Host OS works, and instead of user-mode servers, Guest OS are used (if second type hypervisors). Hypervisors of the first type do not use the services of the host OS and run directly on the hardware. The advantage of exo-kernel-based virtualization is that the display layer is eliminated. Unlike the Virtual Machine Monitor, which must maintain address translation tables on disk (and other tables for converting virtual resources to physical resources), there is no need for such reassignment when using an exo-kernel. The exo-kernel only needs to keep track of which virtual machine (VM) which hardware resources were transferred to.

The exo-kernel architecture does not require creating a copy of the real computer and isolating the virtual machines from physical resources. According to this approach, each virtual machine is provided with a subset of the real resources of the computer. An exo-kernel executed in privileged mode distributes computing resources between virtual machines and controls their use, so that none of the machines tries to use alien computing resources. Virtual machines can serve as software containers that encapsulate the virtual hardware, the Operating System and the applications. Migrating software agents together with a virtual machine is easier than moving individual agents, since when using VM, the local software agent group moves along with the environment required for it (configuration files, system tables, etc.).

The development of MASKRP based on the exo-kernel OS includes the definition of agent roles, messages, delivery environment properties, states, transitions and

actions of the agents, synchronization mechanisms, failure models and other characteristics describing agents and their interactions. The development can be carried out using special instrumental software tools, for example, Multi-Agent-KB5 toolkit [2], which allow knowledge experts to create the problem-oriented MASKRP based on high-level abstractions implemented by high-performance specialized system software modules.

2 Organization of the Problem-Oriented Multi-Agent Solver

An example of a two-level structure of a problem-oriented Multi-Agent Solver of one of the MASKRP nodes is shown in Fig. 1. The functional and structural organization of this solver is described in more detail in paper [2].

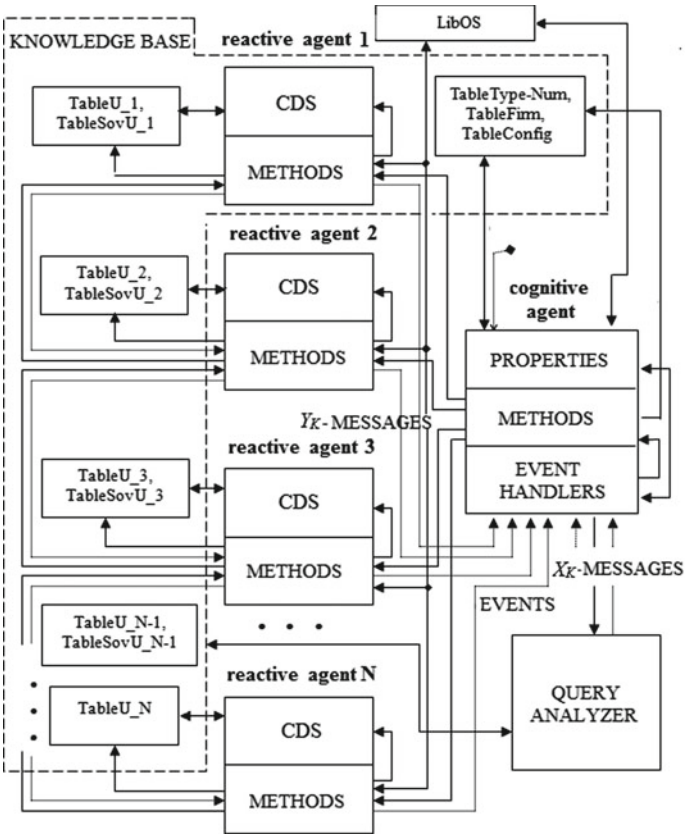


Fig. 1 Structure of the local problem-oriented multi-agent solver

The problem solved at the computing node is divided into separate subtasks, which are distributed among the reactive agents. The work of the reactive agents is coordinated by the cognitive agent. The priorities of the reactive agents included in the group are set according to the sequence number of the program agent. In this case, the first program agent is associated with the highest priority knowledge base tables: Table_1, TableSov_1. The program agent with the number N is associated with the least priority tables: Table_N, TableSov_N. The local Multi-Agent Solver also uses additional knowledge base tables.

For processing knowledge of the problematic components, special Cognitive Data Structures (frame-scripts) are used that describe the plan for solving the target task, define the rules for reducing the task to the subtasks, and establish links between the software agents responsible for solving these subtasks. Four types of methods are implemented for working with the knowledge base: comparison (CMP), association (ASS), analysis (ANS) and specification (VAL). CMP-method is called when comparing events or objects; ASS-method is used to receive responses to requests for relationships between objects and events; ANS-method realizes the logical analysis of events. Fuzzy queries can be used to specify objects in the MASKRP (VAL-method). Fuzzy queries allow to find records in the knowledge base that satisfy the query to a certain extent.

The state and transition diagram of a cognitive software agent is shown in Fig. 2.

The interaction scheme of the computing node reactive agents is shown in Fig. 3.

Each computing node of the MASKRP can use the concurrency model in which several computing threads (agents) are in the execution state, and only one of

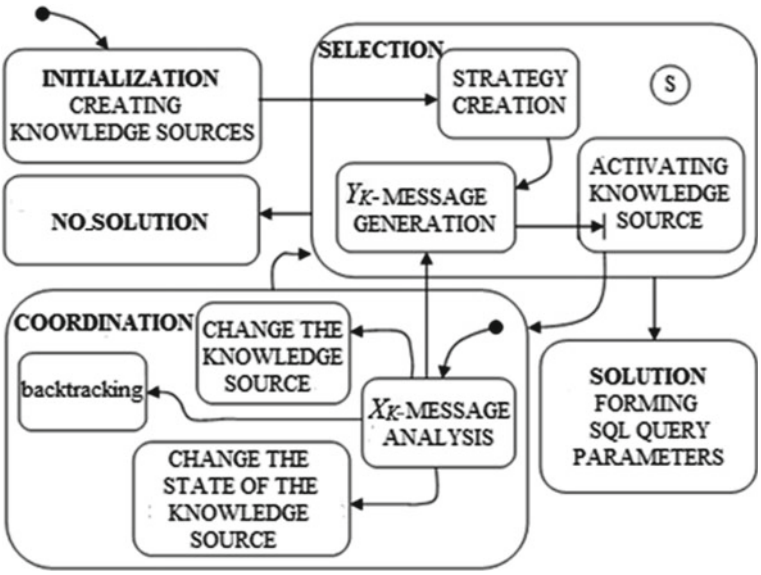


Fig. 2 State and transition diagram

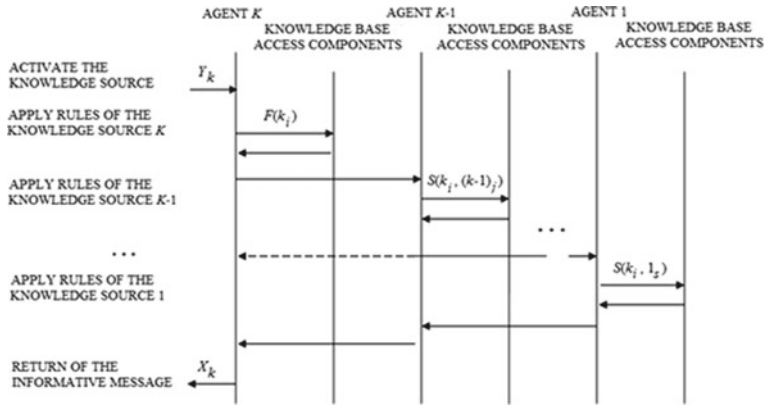


Fig. 3 Interaction scheme of the reactive software agents

the threads is actually executing on the processor at any given time. This concurrency model uses shared memory to exchange information. Competitive threads are described by the consistency model, which defines the order in which operations performed by local agents in a node should be executed, and the order in which the results of these operations should be transmitted to the group members.

In addition to concurrency (competitive) computations, parallel (simultaneous) computations can be implemented in the MASKRP nodes. To implement parallel computing, Uniform Memory Access (UMA) multiprocessors are usually used. In this case, the whole set of software agents is divided into subsets (groups). Agents that belong to different groups can act simultaneously. Agents of the same group consistently coordinate their local solutions and take turns using shared computing resources.

The agents are divided into groups using the compatibility matrix S and the inclusion matrix R. The compatibility matrix S has the following form (1):

$$S = \begin{bmatrix} 0 & s_{12} & s_{13} & \cdots & s_{1M} \\ s_{21} & 0 & s_{23} & \cdots & s_{2M} \\ s_{31} & s_{32} & 0 & \cdots & s_{3M} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_{M1} & s_{M2} & s_{M3} & \cdots & 0 \end{bmatrix} \begin{matrix} S_1 \\ S_2 \\ S_3 \\ \\ S_M \end{matrix} \quad (1)$$

where $s_{ij} = 1$, if the agents A_i and A_j use different computing resources and work in parallel, otherwise $s_{ij} = 0$.

The distribution of agents into groups is based on the inclusion matrix R (2):

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1M} \\ r_{21} & r_{22} & \cdots & r_{2M} \\ \cdots & \cdots & \cdots & \cdots \\ r_{H1} & r_{H2} & \cdots & r_{HM} \end{bmatrix} \begin{matrix} R_1 \\ R_2 \\ \\ R_H \end{matrix} \quad (2)$$

where M is the number of agents, H is the number of groups. $r_{ij} = 1$ if the agent A_i is included in the group Y_j. The agent A_i is included in the group Y_j if $S_i \cap R_j = \emptyset$ that is the matrix rows do not intersect. For optimal partitioning into subsets it is necessary to consider the functional features of the software agents, their requirements for computing resources, as well as know the structural organization of the MASKRP node used to implement parallel computations.

To create a single virtual address space in parallel computing systems that do not use Remote Memory Access (include Clusters of Workstations), Distributed Shared Memory (DSM) and Object-Oriented version of Remote Procedure Call (RPC) are implemented. Standard high-level Message-Oriented Middleware (MOM) is used for communication between MASKRP nodes. MOM allows to organize the interaction of software agents in distributed computing systems consisting of heterogeneous nodes, combined in the form of a virtual organization. The inter-node interaction of software agents in a virtual environment poses a number of problems associated with increasing network load, increasing message delivery time and searching time for the necessary agent. To improve the performance of the Multi-Agent Solver and reduce the total volume of data transferred over the network, the agent-associated distributed knowledge base is divided into clusters. At the same time, two complementary mechanisms for interaction of software agents are used: The mechanism of localization of inter-agent interactions and the mechanism of dynamic distribution of agents in the MASKRP nodes.

3 Optimization of the Logical Structure of the Distributed Knowledge Base

It is advisable to consider the synthesis of the logical structure of a distributed knowledge base as a sequential solution of three particular problems:

(1) determining the optimal localization in a computing system of data groups (knowledge sources) that provides a minimum of total network traffic and satisfies the specified constraints;
(2) determining the structure of the optimal distribution of groups by types of logical records, which provides a minimum of the total time of local information processing at the nodes of the computing system under the specified constraints;
(3) determining the structure of localization of the knowledge base in the nodes of the computing system, which provides the minimum value of the total time of access to the nodes of localization and knowledge processing.

In this case, the logical structure being developed should provide the minimum implementation time for operational requests. To do this, a certain part is allocated from the logical structure, which contains only those types of records whose elements are used in forming output structures on request. The structure of the links between the records of this part of the logical structure should minimize the response time of the system to the request.

The analysis of the properties and combinatorial features of the problem allowed us to develop an effective synthesis algorithm. The developed algorithm is based on the procedures for decomposing the original problem into a number of interrelated subproblems of smaller dimension, solving the selected subproblems using known methods, composing local optimal solutions, and generating the results of solving the general problem.

Let us consider an algorithm for solving the problem of synthesis of the optimal logical structure of the DKB and the structure of localization of the knowledge base by the criterion of the minimum total time for the implementation of a set of user requests. The algorithm consists of three interrelated steps.

Step 1. At this step, the localization of knowledge sources in the computing system is determined by the criterion of the minimum total traffic. To solve this problem, an approximate algorithm for the distribution of DKB clusters between the MASKRP nodes is used. This problem belongs to the class of linear integer programming problems, and as a result of its solution, the optimal localization of knowledge sources on the nodes of the computer network is determined by known methods.

Step 2. This step solves the problems of optimal distribution of data groups of each node by types of logical records according to the criterion of the minimum total time of local data processing in each node of the network.

The number of synthesis tasks for this stage is determined by the number of network nodes. The initial data are the results of the first step that is, the subgraphs of the graph of the canonical structure of the DKB, as well as the time and volume characteristics of the subgraphs of the canonical structure of the DKB, the set of requests of users and network nodes [13–15].

The following restrictions are used:

- limits on the number of groups in a logical record,
- on the one-time inclusion of groups in records,
- on the cost of storing information,
- to the required level of information security of the system,
- for the duration of operational transactions on distributed knowledge base servers,
- for the total service time of operational requests on the servers.

As a result of solving the considered problems, the logical structures of the knowledge base of each node of the computer network are determined.

Step 3. This step solves the problem of localization of the knowledge base on the nodes of the computer network. As the initial data for its solution, the results obtained in the first two steps, as well as the characteristics of the DKB, are used. This problem is solved using the following constraints:

- on the total number of synthesized logical records placed on the MASKRP nodes;
- the amount of available external memory of the network servers for storing the knowledge base;
- the number of copies of logical records placed in the MASKRP.

As a result of the proposed algorithm, matrices of localization of a set of data groups by logical record types (the result of step 1) and further groups of records by MASKRP nodes (the result of step 2) are formed [14]. The result of the last stage is a matrix of localization of groups of logical records on nodes of the Multi-Agent System of knowledge representation and processing.

4 Conclusion

The article discusses the issues of improving the performance of a Multi-Agent System of knowledge representation and processing. The structure of the Multi-Agent Solver with two types of software agents is described. The diagrams of states, transitions, and interactions of a group of software agents of the local MASKRP node are presented. The methods of knowledge processing used in the MASKRP and the types of requests to the DKB are considered.

An approach to improving the performance of the MASKRP is proposed, which includes the efficient distribution of knowledge sources in the Multi-Agent Solver of the computing node, as well as the construction of an optimal logical structure of the Distributed Knowledge Base. An approach to the implementation of MASKRP management subsystems based on an exo-kernel operating system and specialized event-driven software modules is considered.

Further work in the framework of this topic is to develop software that implements algorithms for searching for variants of the logical structure of the DKB, which provides the optimal value of a given criterion for the effectiveness of functioning and satisfies the main system, network and structural constraints.

References

1. Zaytsev EI (2008) Method of date representation and processing in the distributed intelligence information systems. Autom Mod Technol 1:29–34
2. Zaytsev EI, Khalabiya RF, Stepanova IV, Bunina LV (202) Multi-agent system of knowledge representation and processing. In: Proceedings of the 4th international scientific conference intelligent information technologies for industry (IITI'19), Springer, pp 131–141
3. Narin'yani AS (1994) NE-faktory i inzheneriya znanij: ot naivnoj formalizacii k estestvennoj pragmatike. Nauchnye trudy Nacional'noj konferencii po iskusstvennomu intellektu. T.1.–Tver': AII:pp 9–18
4. Alouf S, Huet F, Nain P (2002) Forwarders vs. centralized server: an evaluation of two approaches for locating mobile agents. Perform Eval 49:299–319

5. Aly S, Badoor H (2018) Performance evaluation of a multi-agent system using fuzzy model. In: 1st international workshop on deep and representation learning (IWDRL), Cairo, 2018, pp 175–189
6. Batouma N, Sourrouille J (2011) Dynamic adaption of resource aware distributed applications. Int J Grid Distrib Comput 4(2):25–42
7. Cardoso R, Hübner J, Bordini R (2013) Benchmarking communication in actor-and agent-based languages. Engineering multi-agent system. Springer, Berlin, pp 58–77
8. Darweesh S, Shehata H (2018) Performance evaluation of a multi-agent system using fuzzy model. In: 1st international workshop on deep and representation learning (IWDRL), 2018, pp 7–12
9. Hadzibeganovic T, Cui P, Wu Z (2019) Nonconformity of cooperators promotes the emergence of pure altruism in tag-based multi-agent networked systems. Knowl-Based Syst 171:1–24
10. Jurasovic K, Jezic G, Kusek M (2006) Performance analysis of multi-agent systems. Int Trans Syst Sci Appl 4:601–608
11. Sethia P (2011) High performance multi-agent system based simulations. India: center for data engineering international institute of information technology
12. Houhamdi Z, Athamena B, Abuzaineddin R, Muhairat M (2019) A multi-agent system for course timetable generation. TEM J 8:211–221
13. Nurmatova E, Shablovsky A (2015) The development of cross-platform applications semistructured data. Herald of MSTU MIREA, 3, 2015
14. Nurmatova EV, Gusev VV, Kotliar VV (2018) Analysis of the features of the optimal logical structure of distributed databases. In: Collection of works the 8th international conference "distributed computing and grid-technologies in science and education. Dubna, 167p
15. Demenkova TA, Tomashevskaya VS, Shirinkin IS (2018) Mobile applications for tasks of distance learning. Russ Technol J 6(1):5–19 (In Russ.)

Cyber-Attack in ICT Cloud Computing System

Pranjal Chowdhury, Sourav Paul, Rahul Rudra, and Rajdeep Ghosh

Abstract Nowadays, cloud system is laid low with cyber-attack that underpin a great deal of today's social group options and monetary development. To grasp the motives behind cyber-attacks, we glance at cyber-attacks as societal events related to social, economic, cultural and political factors (Kumar and Carley in 2016 IEEE conference on intelligence and security information science (ISI). IEEE, 2016 [1]). To seek out factors that encourage unsafe cyber activities, we have a tendency to build a network of mixture country to country cyber-attacks, and compare the network with different country-to-country networks. During this paper, we gift a completely unique approach to find cyber-attacks by exploitation of three attacks (Zero, Hybrid and Fault) which may simply break the cyber system. In hybrid attack, we use two attacks, dictionary attack and brute-pressure attack. Firstly, we analyze the system and then lunch the attacks. We have a tendency to observe that higher corruption and an oversized Web information measure favor attacks origination. We have a tendency to additionally notice that countries with higher per-capita-GDP and higher info and communication technologies (ICT) infrastructure are targeted a lot often (Kumar and Carley in 2016 IEEE conference on intelligence and security information science (ISI). IEEE, 2016 [1]).

Keywords Fault attack · Hybrid attack · Dictionary attack · Brute-pressure attack · Zero-day attack

1 Introduction

Information and communication technology (ICT) structures underpin a lot of today's societal features and financial development. Cyber-attack affects economically and socially [2]. In cloud computer system, there are many Websites attacked by many types of cyber-attack like fault attack, dictionary attack, brute-pressure attack and hybrid attack etc., which are making a bad effect economically. Right now, e-banking

P. Chowdhury (✉) · S. Paul · R. Rudra · R. Ghosh
Institute of Engineering and Management, Sector V, Kolkata, West Bengal 700091, India
e-mail: pranjal.prim@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_12

system also is not secure for cyber-attack. Because hacker can easily break the firewall of this e-banking system, and they steal a lot of data and money which are making the economy down. The hackers also hack the social site and publish many false news in the social media and make collations. So, we should prevent this type of problems and became careful and make strong security system in our device and social site.

Cyber-attack is a common thing in every country. Nowadays, every social site and banking system is not secure for cyber-attack. From many banking systems, they get many data which are very important. They can easily can know the bank details, Aadhar number, date of birth, PAN number, and they can easily track all details of that person; sometimes they blackmail them and demand a lot of money. They get their retina and finger print. By using these, they do many illegal works. They hack their social site and use them for a bad mentality which affected manpower system. It is also affected in political system also. First, when hacker hack any computer or any site, they first hack the hard disk. If we catch them, we need to find their IP address. We can do it by key lockers. Key locker means when someone hack something, he/she put off some abidance. The security system is covered by firewall protector. Firewall means the close door of computer security. Hacker break the firewall, then enter the system. When they hack the security system, alert activates. But in our lacking security, we cannot get that alert system. Due to that reason, they can easily enter the system. In next, we discuss some attacks and process of attack, how they break the system firewall and enter the system.

In next step, we discuss different type of attacks, a hacking system procedure, result and conclusion. In second section, we discuss three attacks; in third section, we discuss a hacking system procedure and its result; in fourth section, we discuss the conclusion and prevention.

2 Attacks

2.1 Fault Attack

Fault attack is a technique where testers or managers use their past experience and knowledge of typical defect in a system and predict or guess the fault. The test cases are particularly created based on prediction to uncover the fault. Fault attack is basically an approach that it can call as a way of testing the product when it is already aware of what effect it is looking for when compared this particular approach with generic testing process. In generic testing process, it is not sure; it is not aware of what type of defect or what specific defect it is looking for. So, it generally understands the requirement and overview of the product features and functionality and create various test cases in order to find defects. So, if it remembers the basic objective of testing to find effects, but when it comes to fault attack, it is basically when it is known, there are this type of defects which is generally find when we talk about such applications. Its test cases are driven in such a way that it wants to uncover this defect

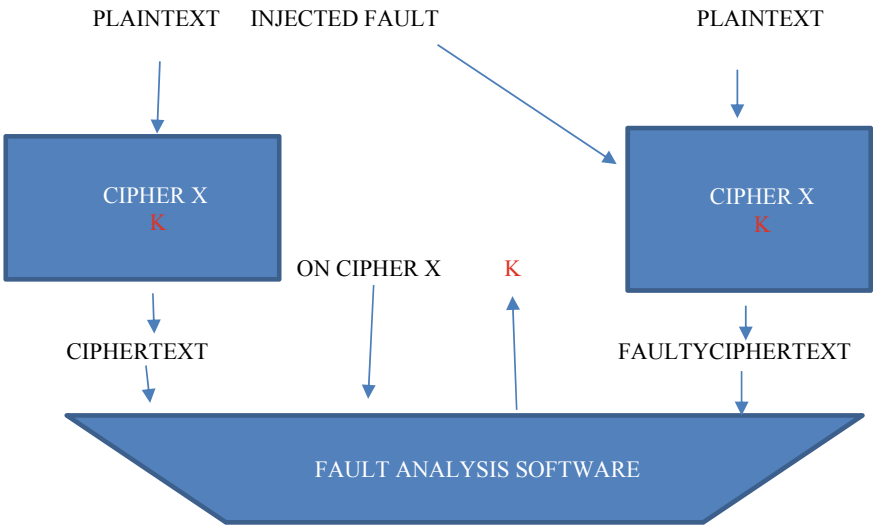


Fig. 1 Fault attack targets: Fault observation and key extraction [4]

from the application, and that is the difference between a generic testing process and the fault attack. The fault attack requires it to have the past experience of similar type of applications or typical types of defects which it gets in a particular product, or even with having a domain knowledge can help us to apply the fault attack approach. Now, this fault attack will generally help us to create certain test cases which are specifically created for particular type of defect, and that is the reason it is said that we are trying to attack the defect with our specific test cases which are specially designed for the same reason and that is called fault attack. It is attacking the fault which is known to us well. Fault attacks described in the family of cryptographic attack mostly apply to cryptographic algorithms, yet such attacks may have an impact on the whole system in a smart card. In this paper, we describe what can be achieved nowadays by using fault attacks in a smart card environment [3] (Fig. 1).

2.2 Hybrid Attack

Hybrid attacks are a form of cyber-attack, wherein the offender blends or extra types of equipment to perform the assault. A common hybrid attack is one which merges a dictionary attack and a brute-pressure attack. The former might include a listing of doubtlessly regarded credential matches. The latter might practice a brute-pressure attack upon every feasible fit. Using a listing of breached usernames and passwords received from earlier incidents, hackers release a credential-stuffing attack conceived of as an easy fit in opposition to the login fields at the goal enterprise(s). However, the attack is amplified via using a brute-pressure technique of attempting extra combos of

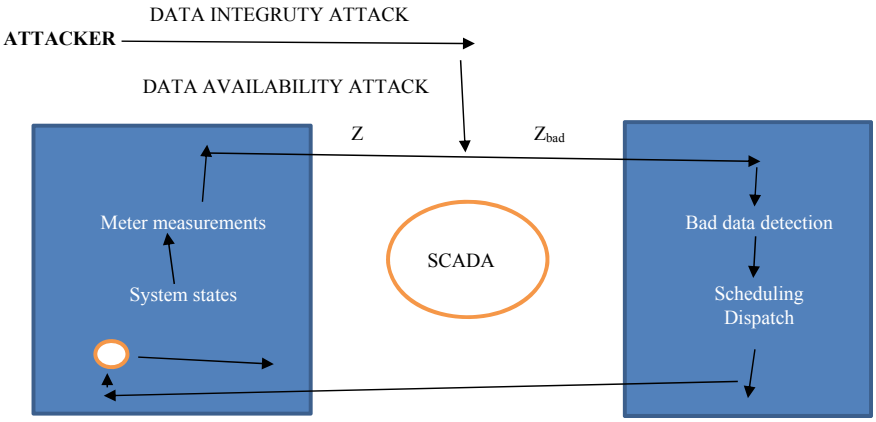


Fig. 2 The schematic diagram of the hybrid cyber-attack [6]

what is understood with the aid of using automating additions of numbers, symbols, and so on in reaction to person behavior to minimally adjust passwords over time [5]. We propose a hybrid attack model and apply typical state estimation processes to check cyber-attacks on power grids during this paper. We propose a hybrid attack model and apply typical state estimation processes to visualize cyber-attacks on power grids throughout this paper. The projected model considers each integrity attack and handiness attack at the same time. Compared with the actual attack, namely, false information injected attack, we tend to analyze their consequences to power systems within the events of false negatives attack and warning attack. The results show that the hybrid attack will confuse the center by manipulating the integrity and handiness of measurements. A lot of significantly, we tend to judge the hybrid attack with totally different values of the value quantitative relation between integrity and handiness attacks, so verify that the hybrid attack can do identical goal with a reduced value [6] (Fig. 2).

2.3 Zero Day Attack

Zero-day attack is that the term accustomed to describe the threat of associate degree unknown security vulnerability during a laptop package or application; for that, either the patch has not been discharged or the applying developers were unaware of or failed to have sufficient time to address. Since the vulnerability is not illustrious in advance, the exploits usually occur while not the data go the users. A zero-day flaw is taken into account as a vital part once planning an application to be economical and secure. Zero-day attack is random attack which cannot be removed. One can only identify it and thus can be avoided; it is also called one-day attack, and it is a threat that tries to exploit computer application and vulnerabilities. The zero-day attack in

networks exploits associate degree undiscovered vulnerability so as to affect/damage network or programs. The term "zero-day" refers to the quantity of days offered to the code or hardware marketer to issue a patch for this new vulnerability [7].

- Keep software up to date to assist defend against a zero-day vulnerability.
- Check for an answer once a zero-day vulnerability is announced. Most package vendors work quickly to patch a security vulnerability.
- Do not underestimate the threat. Cybercriminals will obtain to take advantage of security holes and gain access to devices and private data. They will use information for a spread of cybercrimes as well as identity theft, bank fraud and ransomware.
- Invariably use a reliable security software to help keep devices safe and secure.

3 Algorithm and Results

The stepwise procedure for cyber-attack is described below.

Step 1: Starting Applet
Step 2: Accessing backdoor convert channel. (System counter Measures)
Step 3: Stepping Botnet (Emanation Analysis)
Step 4: Launching Fault Attack
Step 5: ACK Piggybacking
Step 6: Final Hybrid attack
Step 7: Launching Fault Attack
Step 8: Continue Evasive man oeuvre
Step 9: Kernel Panic
Step 10: System counter Measures
Step 11: Backdoor access successfully
Step 12: User Authorized

First, we are starting applet. Then, starting system counter measures. We are accessing backdoor convert channel. We do emanation analysis and step botnet. After that, we are launching fault attack. The tracking IP is started. Then, we do ACK piggybacking. Finally, we lunch hybrid attack. Again, we launch fault attack and continue evasive maneuver. The kernel panic continues and the system counter measures. The backdoor access successfully and user authorization done. This system was protected by multilayer firewall system. For breaking this firewall, we use two attacks. Hybrid and fault attack. By using these attacks, most of the firewall can easily break and enter the system. Hackers are mainly use these types of attacks. Fault attack and hybrid attack successful work about 70%. In this project, first we analysis the system and lunching the fault attack. Then, ACK is piggybacking and lunching hybrid attack. The firewall is broken down and backdoor access successfully. User can easily enter this system.

4 Conclusion

In this analysis, we have a tendency to check out the cyber-attacks as a social phenomenon. We have a tendency to used network analysis to grasp the motivations behind the attacks. Victimization network analysis of attacks, we discover India and the United States of America attack the most, which the USA could be attacked the most. The analysis highlights that there is a medium correlation between information measure of cyber-attacks and network bandwidth of a country [8]. This correlation indicates that prime bandwidth countries are smart supply of bots for creating cyber-attacks, probably as a result of high bandwidth countries may facilitate to host computers that may execute high bandwidth cyber-attacks. This research is a starting in understanding the advanced network of country-to-country cyber-attacks [8]. We have a tendency to should alert all time regarding our device system. We have a tendency to must use two-step verifications in our social website for that hacker cannot enter the system simply. Some social sites like Facebook WhatsApp currently use multi-steps verification system within the website so that hackers can-not enter easily into the system.

5 Future Degree of Research

Cyber-attack is a common phenomenon in every country. Cyber-attack is occurring in many sectors like banking system, social sites etc. The attackers can easily know the bank details, Aadhar number, date of birth, pan number, and they can easily track all details of that person; sometimes, they blackmail them and demand a lot of money. They get their retina and finger prints.

This research paper has great future degree of research. The algorithm which is used in this research paper can be used in credit card transaction to ensure safe transaction of money. It can also be used to protect against unauthorized access to government data. It will also help in evading loss of important like, Aadhar number, date of birth, pan number, retina scan etc.

References

1. Kumar S, Carley KM (2016) Approaches to understanding the motivations behind cyber-attacks. In: 2016 IEEE conference on intelligence and security information science (ISI), IEEE
2. Kiss I, Genge B, Haller P (2015) A clustering-based approach to get cyber attacks in methodology management systems. In: 2015 IEEE 13th international conference on industrial IP (INDIN), IEEE
3. Giraud C, Thiebeauld H (2004) A survey on fault attacks. Consumer credit analysis and advanced applications VI. Springer, Boston, pp 159–176
4. Sheikh R, Cammarota R (2018) Improving performance and mitigating fault attacks victimization worth prediction. Crypt 2(4):27

5. https://www.hypr.com/hybrid-attack/
6. Tu H et al (2020) A hybrid cyber-attack model for cyber-physical power systems. IEEE Access 8:114876–114883
7. Al-Rushdan H et al (2019) Zero-day attack detection and hindrance in software-defined networks. In: 2019 International Arab conference on data technology (ACIT). IEEE
8. Kumar S, Carley KM (2016) Approaches to understanding the motivations behind cyber-attacks. In: 2016 IEEE conference on intelligence and security scientific discipline (ISI), IEEE, pp 307–309

Internet of Things: A Review on Its Applications

C. A. Irfana Parveen, O. Anjali, and R. Sunder

Abstract The Internet of Things (IoT) is the ambassador of computing technology and is the next phase in evolution of Internet. It is the interconnection of billions of smart things that contain objects with sensors, actuators, networking and technology which work together to make this "network of networks of autonomous objects" a reality by providing smart services to the end-users. In this paper, we give an overview about IoT and its history. Secondly, we discuss some of the applications of IoT mainly smart home, smart health care, smart farming, smart city and smart industry as well as blockchain application in industry. Finally, the issues and challenges are analysed to encourage more investigations into the domains.

Keywords IoT · Smart home · Smart health care · Smart farming · Smart city · Smart industry · Blockchain

1 Introduction

Internet of Things or IoT is the network of smart objects or things which can exchange data between them and can transmit (or receive) information with other devices and systems. It is one of the crucial factors in digital transformation of cities and industries. IoT is an umbrella term as these smart devices are incorporated with sensors, actuators, networking, software and other technology. IoT devices range from common household items like smart lights to advanced industrial tools. Sensors or devices, connectivity or network, data processing and a user interface are the core components of the Internet of Things. Here, the data is collected by sensors which

C. A. Irfana Parveen (✉) · R. Sunder
Sahrdaya College of Engineering and Technology, Thrissur, Kerala, India
e-mail: irfanarahi14@gmail.com

R. Sunder
e-mail: sunder@sahrdaya.ac.in

O. Anjali
Rajadhani Institute of Science and Technology, Palakkad, Kerala, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_13

are sent to the cloud. These sensors are connected to the cloud by Wi-Fi, Bluetooth, cellular or direct Internet connection via Ethernet. Once these data reach the cloud, the software processes it, and then it is moved to the end-users [1].

Despite the fact that the Internet of Things concept was initially proposed almost two decades ago, the technologies that support IoT had previously existed and had been under development for several years. Main technologies of IoT, like Advanced Research Project Agency Network (ARPANET), were introduced in 1969, radio frequency identification (RFID) in 1973, beginning of TCP/IP in 1974, Domain Name System (DNS) in 1984 and World Wide Web (WWW) in 1989 [2].

In 1990, the Internet was widely used in both commercial and consumer industries, and ubiquitous computing was proposed in 1991. Google was incorporated in 1998. And 1999 was the big year for IoT as the term Internet of Things was coined by Kevin Ashton. From 2000 onwards due to digitalization, the Internet plays a significant role in business and production. Finally, during 2008 to 2009, the Internet of Things was "Born" [2, 3]. According to the CISCO analysis, only 500 million interactive devices in 2003 and by 2025 it will be more than 75.44 billion connected devices [4].

2 IoT Applications

A review of chosen IoT applications is presented in this section. Existing areas of interest for applications such as smart home, smart health care, smart farming, smart city and smart industry are depicted in Fig. 1.

2.1 Smart Home

Smart homes are a range of smart devices with Internet access. It is also referred to as home automation which helps to monitor and control all processes taking at home and provide better security and comfort to the homeowners. By studying the energy consumption of each device, we can save energy and cost by reducing their usage and shutting down when not in use. IoT allows these smart devices to automate, connect and communicate with each other without your participation. Some of the smart home technologies are smart TV, smart light, smart thermostat, smart door lock, smart security cameras and smart parking [5].

To build a strong infrastructure for a smart home, we need to select between wired and wireless networks. If one device is only used like a video doorbell or smart thermostat, we can connect these to WI-FI. As devices increase, we can add more routers to cover the same. Wireless set-ups are mainly used for lighting and security. A wired system can connect these devices to the Internet without interruption [6].

One of the fundamental aspects of the smart home is the smart door lock system. Smart lock is an electromechanical lock which automatically locks and unlocks the door without physical interaction. It can connect to the network and can be controlled

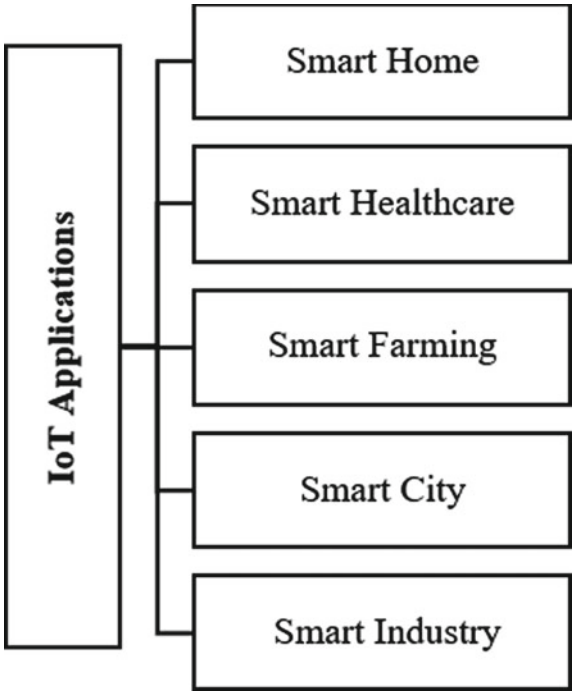


Fig. 1 Applications of IoT

by smartphone. The primary goal of this smart device is to keep unauthorized people out of the residence [7]. Authorized devices which give instructions to these smart locks use cryptographic keys and wireless protocol to execute the process [8]. Some of the features of the smart door lock are that they can record if someone locks or unlocks, they can report if the lock is tampered, and they can connect to various services and security [9]. Keypad and combination keys, fingerprint and biometric keys are few types of smart door lock available.

Smart home application provides smart energy management (SEM) to monitor energy consumption. The main objective is to save energy and reduce the cost with effective operation and planning. The SEM system maximizes the use of devices which require more energy during the hours when energy cost is low [10]. SEM is based on the NodeMCU and Android platforms. They can monitor, optimize and measure energy consumption [10]. It can also control the device from anywhere using remote control. It sends notification when something fails due to power cut-off, overcurrent, short circuit or over temperature [10].

Smart homes can also support elderly people living alone with the help of IoT [11]. The system is designed according to which it detects patient's activities by using sensors. These activities can be normal and abnormal. When abnormal activities are detected, the patients are reminded that something went wrong. During difficult

situations, the patient's relatives or caregivers are contacted using a wristband which is equipped with an emergency button [11]. The system provides privacy and ensures that it tracks the routine without any intrusion.

2.2 Smart Health Care

Health care, often known as the Internet of Medical Things (IoMT), is one of the fastest growing sectors in the IoT market. It is one of the most important aspects of a nation's development and represents the society's overall well-being. The biggest challenge faced by the healthcare system in recent times is due to the increase in population, and the medical-related problems also increase. Before IoT, patient's interaction with doctors was limited, and there was no way a doctor or hospital could monitor the patient's health continuously and recommend accordingly. In the healthcare sector, IoT-enabled devices make remote monitoring on patients making them healthy and safe and also enable physicians to deliver proper care. This results in reducing length of hospital stay, prevents readmissions, increases the number of positive healthcare outcomes, reduces healthcare cost, improves treatment and so on [12].

In IoT health care, the patient's data are collected by using sensors, and these data are processed using artificial intelligence algorithms. Then the data are subsequently moved to the cloud. According to that, the doctors and health practitioners take action upon this information [13]. Some of the healthcare IoT devices are insulin pen, smart watch, connected contact lens, smart video pills, brain swelling sensors, connected inhalers and ingestible sensors.

IoT health care plays a vital role in fighting COVID-19. The main issue during the pandemic is unreachability to the patients which can be resolved with the help of Internet of Healthcare Things (IoHT) or Internet of Medical Things (IoMT) [14]. These smart devices help people who are in quarantine by monitoring and measuring their blood pressure, glucose level and heartbeat, thereby reducing the workload of medical staff and connecting to service providers during emergency situations [14]. The most advantageous thing in IoT health care is without human interaction the information about patients will be collected by these devices. There are smartphone applications which tell the user how close they are to a corona positive person, and they can be more cautious before stepping outside the house [14].

Several contributions are made by IoT health care, and some of the major researches are described in subsequent paragraphs.

Kaur and Jasuja [15] propose a system which monitors body temperature and pulse rate of a person with the help of temperature sensors, heart rate sensor and Raspberry Pi device. Muhammad et al. [16, 17] proposed a voice disorder assessment and treatment system with the help of temperature sensors, ambient humidity sensors, electrocardiogram, microphone and smartphone.

Satija et al. [18] proposed a system for monitoring ECG with signal quality assessment with the help of ECG sensors and Arduino and Android phone device. Hegde

et al. [19] introduced a system for COVID-19 pre-screening, fever and cyanosis non-contact detection. Here temperature estimates using infrared camera image and cyanosis from lip images. Using PoseNet, they detect and segment the forehead and lip region. With the help of a deep learning model, Priyadarshini et al. [20] introduce a system for predicting stress types, diabetes and hypertension attacks.

2.3 Smart Farming

As the population increases, the agricultural sector must satisfy the demands in any climate or weather conditions. To achieve that, it needs to adopt new technologies like IoT, GPS and big data. This results in smart agriculture and precision farming which increase the quality of products and reduce the cost and waste.

In IoT-based smart farming, the sensor is used in every step of farming to monitor the crop yield from its seed to a fully grown plant. It gives relevant information like temperature, moisture, sunlight, humidity, water, soil and also how much time and resources are required for the same. It also captures the live data, and it is sent to appropriate users so we do not require to observe the field occasionally. Once we get the information, we can examine it and can be utilized to prevent the loss of yield in future by recording the data [21]. Smart farming makes farmers well organized, intelligent and connected. They will have a significant amount of information about crop yields, machinery, climatic factors, fertilizer, soil condition, water and so on.

Precision agriculture is a farming concept based on observation and measurement that utilizes contemporary data apparatus and ensures that the harvest and soil take place in a beautiful way to get a most positive creation. As IoT is implanted with electronic sensors and programming, it permits these gadgets to send and get information from one another. Also, they store, detect, gather and interface these gadgets to the web.

IoT also focuses on waste management by improving waste collection and recycling [21]. Route optimization minimizes fuel costs and human labour while emptying dumpsters throughout the city. These trash containers contain ultrasonic sensors which are set on the highest level to identify the trash level, so when it reaches above 90%, it will send information to the advisory group by telling the trash had arrived at its extreme level [21]. After this, the dustbin should gather as quickly as time permits. A server is utilized to store information about the dustbin levels [21].

IoT applications in agriculture, like precision farming, crop management, greenhouse automation, waste management, enhance the product quality, lessen human labour and expense, increase business efficiency and help in cost management and waste reduction [22].

2.4 Smart City

The Internet of Things (IoT) lies at the heart of smart city development. More than 50% of the world's population have been living in cities since 2007, and it will increase to 60% by 2030 [23]. As population increases, demand for resources also increases. They need basic necessities like water, energy, environmentally friendly transportation, health care, clean air and waste management [24]. So by combining IoT technology with smart cities, these demands can be achieved effectively and can provide better quality of life and services to the residents. Some of the core infrastructure elements in smart cities are adequate water supply, health and education, sustainable environment, assured electricity supply and so on [25].

Smart infrastructure is essential for building smart cities. It is mostly composed of physical components such as sensors, actuators, machines and cabling, as well as fundamental ideas such as data, analytics, feedback and flexibility [26].

An intelligent parking system (IPS) is one of the most important aspects of a smart city. It is an advanced automatic driving system with two functions like parking guidance and damage notification [27]. It can provide easy parking; it can park at suitable spots without delay, and it gives notification to the user if the car got damaged [27].

Smart city also played a main role in global pandemic like COVID-19. Here in the smart city network, the system recognizes people who do not wear mask in public areas with the help of closed-circuit television (CCTV) cameras [28]. If the person is detected without a mask, the information is passed to the corresponding authority, and necessary actions are taken. So, this implementation impacts on reducing the spread of disease [28].

Smart city based on IoT is one of the key research topics as it is the integration of diverse algorithms like IoT, artificial intelligence, big data, security, image processing, software architecture and so on [29]. Artificial intelligent algorithms are used to identify and automate elements like people, vehicles, buildings, etc. Security algorithms are used to manage the data, also it provides privacy and eliminates risk within the smart cities. Software architectures help to integrate the data of IoT devices [29]. In the following paragraphs, we will discuss some most important smart city studies.

As cloud environments are weak to maintain data integrity, Kubendiran et al. [30] implement a blockchain technology which provides stable and secure technology in e-health systems and also introduced an algorithm which has a shorter data discovery time. As there are some issues for cloud computing in smart city, Choi and Ahn [31] suggest a fog computing paradigm as fog nodes are close to IoT devices. They also support various requirements like storage, CPU, delays and scalability.

Park et al. [32] introduced the smart home framework called S-mote for household applications in the IoT network. Here if we add the infrared or radio frequency module to the home appliances, we can operate all IoT functions using a single application. To safeguard secret digital documents, Halder et al. [33] proposed steganography protocol. It can hide the entire content, so unauthorized access to this document is not possible. Ding et al. [34] proposed a precise forecast index system. The model

was built using grey relational analysis and combination of backpropagation neural networks and support vector machines. It gives more accuracy than the existing system.

2.5 Smart Industry

IoT is the basis for industrial transformation which is also referred as Industrial Internet of Things (IIoT) or Industry 4.0. It is the connection of machines, computers, tools, sensors and people which give a new approach to production and reshape the industrial operations to a business outcome. It ensures that all company processes are controlled by a central system (automated management), it detects whether a machine requires any maintenance before production (predictive maintenance), at every stage it monitors the quality of products (quality control), and it can obtain real time in transit information at increasing efficiency (supply chain optimization) [35].

Even though IIoT is gaining popularity in academia and industry, it raises a number of privacy and security concerns. As a result, to ensure the IoT security, we introduce a technology called blockchain [36]. It is also referred to as distributed ledger technology (DLT) which utilizes decentralization and advanced encryption algorithms so that the data can be unalterable and transparent, and thereby we can send and receive secure messages among connected things. The core components of blockchain are node, transaction, block, chain, miners and consensus [35].

Blockchain application in some of the industrial sectors is shown in Fig. 2 and summarized below:

(i) Healthcare industry: Blockchain-enabled IoT can be used for protecting healthcare data and managing electronic medical record data [37].
(ii) Supply chain or logistics: Blockchain technology improves the customer experience by enabling transparency, tracking and enhancing licence of goods, services and software [38].
(iii) Power industry: Blockchain technology reduces cost, provides a sustainable environment and enhances privacy and security in the energy sector [39].
(iv) Agriculture industry: Blockchain technology optimizes the food supply chain, provides peer-to-peer agricultural trade and processes complex agricultural deals [40].
(v) Manufacturing industry: Blockchain technology assures quality, enhances track and trace and also protects intellectual property [41, 42].

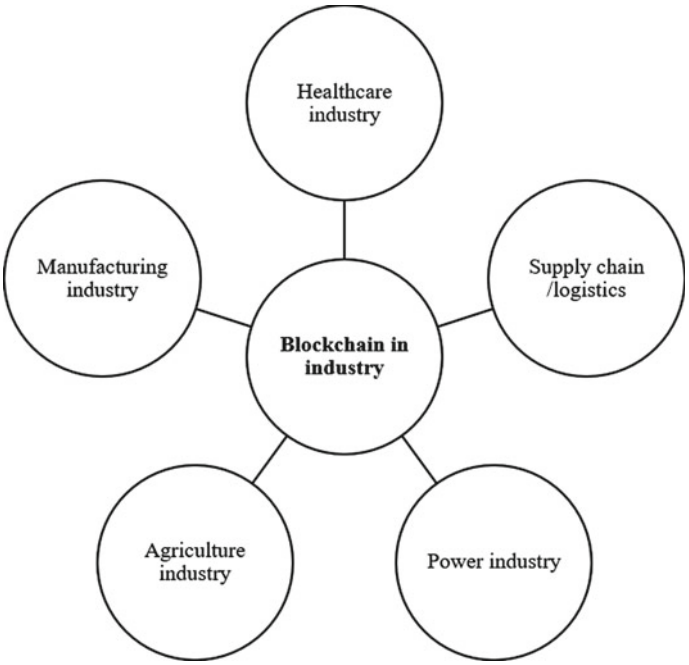


Fig. 2 Blockchain application in industry

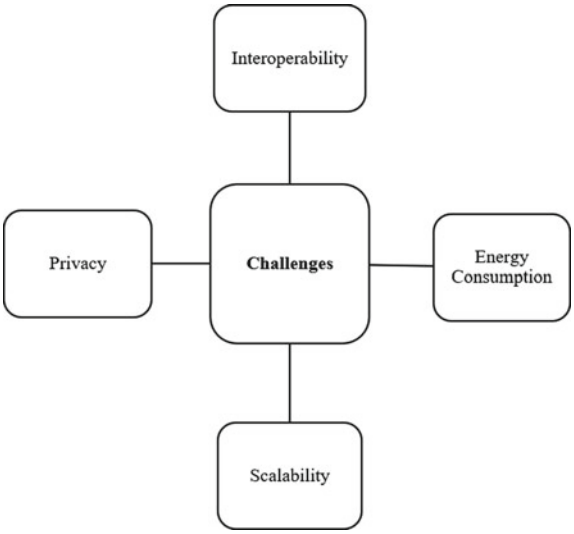


Fig. 3 Challenges in IoT

3 Issues and Challenges

Due to data transmission and storage, there exist many issues and challenges faced by IoT, and some of them are depicted in Fig. 3. As a result, before implementing the concept in current systems, these issues must be examined further:

- The exponential growth of connected devices is a concern as it increases energy consumption [43].
- Lack of interoperability means data cannot be effectively exchanged between overlapping connected devices and systems. Still 60% of IoT devices are locked due to lack of standardization [44].
- As the idea of IoT increases, there is a big challenge for "scalability of IoT". These IoT devices are not effectively adapting the changes in the environment and meet the changing needs in future.
- Despite the fact that IoT devices provide users with a good experience, they are always surrounded by some security issues. Insecure connections and data storage are the most common security concerns in IoT applications [45].
- IoT devices have low storage capacity as the information is stored in the nodes.
- Many IoT devices run low-quality software to some kinds of vulnerabilities.
- The other challenge faced by IoT devices is the complex nature as its development cycle is longer and more intensive [46].

4 Conclusion

The IoT is advancing our life, business and work by offering some pretty applications like smart home, smart health care, smart farming, smart city and smart industry. In this paper, firstly we focused on the background of IoT and its brief history. Then we further discussed how IoT is being used in various fields, particularly the current research trends and also about blockchain technology in various industrial sectors and finally the challenges and issues faced by the connected devices.

In future, we will focus on specific topic like blockchain technology in health care only and take effort to offer all the IoT's benefits to end customers via their smartphones.

References

1. IoT explained—how does an IoT system actually work? (2021) In: Leverege.com. https://www.leverege.com/blogpost/iot-explained-how-does-an-iot-system-actually-work. Accessed 14 Aug 2021
2. Internet of Things (IoT) history | postscapes (2021) In: Postscapes. https://www.postscapes.com/iot-history/. Accessed 14 Aug 2021

3. Balas V, Solanki V, Kumar R, Khari M (2018) Internet of things and big data analytics for smart generation. 33
4. Evans D (2011) The internet of things: how the next evolution of the internet is changing everything, CISCO. Int J Internet 3(2):123–132
5. What is smart home or building (home automation or domotics)? Definition from WhatIs.com (2021) In: IoT agenda. https://internetofthingsagenda.techtarget.com/definition/smart-home-or-building. Accessed 14 Aug 2021
6. A strong infrastructure for your smart home design (2021) In: Bravas.com. https://bravas.com/bravas-blog/item/a-strong-infrastructure-for-your-smart-home-design. Accessed 14 Aug 2021
7. Mistry I, Tanwar S, Tyagi S, Kumar N (2020) Blockchain for 5G-enabled IoT for industrial automation: a systematic review, solutions, and challenges. Mech Syst Signal Process 135:106382
8. Smart lock—Wikipedia (2021) In: En.wikipedia.org. https://en.wikipedia.org/wiki/Smart-lock. Accessed 14 Aug 2021
9. What is a smart door lock? Access control for security systems—DIY smart home solutions (2021) In: DIY smart home solutions. https://www.diysmarthomesolutions.com/smart-door-lock-access-control-security-systems/. Accessed 14 Aug 2021
10. Tastan M (2019) Internet of things based smart energy management for smart home. KSII Trans Internet Inf Syst. https://doi.org/10.3837/tiis.2019.06.001
11. Sokullua R, Alper Akkas M, Demir E (2020) IoT supported smart home for the elderly. Internet Things 11:100239
12. IoT in healthcare industry | IoT applications in healthcare—Wipro (2021) In: Wipro.com. https://www.wipro.com/business-process/what-can-iot-do-for-healthcare-/. Accessed 14 Aug 2021
13. IoT in healthcare: benefits, use cases, challenges [guide] (2021) In: Intellectsoft blog. https://www.intellectsoft.net/blog/iot-in-healthcare/. Accessed 14 Aug 2021
14. Singh R, Javaid M, Haleem A, Suman R (2020) Internet of things (IoT) applications to fight against COVID-19 pandemic. Diabetes Metab Syndr Clin Res Rev 14:521–524. https://doi.org/10.1016/j.dsx.2020.04.041
15. Kaur A, Jasuja A (2017) Health monitoring based on IoT using raspberry PI. In: 2017 international conference on computing, communication and automation (ICCCA), Greater Noida, pp 1335–1340
16. Muhammad G, Rahman SMM, Alelaiwi A, Alamri A (2017) Smart health solution integrating IoT and cloud: a case study of voice pathology monitoring. IEEE Commun Mag 55(1):69–73
17. Muhammad G, Alhamid MF, Alsulaiman M, Gupta B (2018) Edge computing with cloud for voice disorder assessment and treatment. IEEE Commun Mag 56(4):60–65
18. Satija U, Ramkumar B, Manikandan MS (2017) Real-time signal quality-aware ECG telemetry system for IoT-based health care monitoring. IEEE Internet Things J 4(3):815–823
19. Hegde C, Suresha PB, Zelko J, Jiang Z, Kamaleswaran R, Reyna MA, Clifford GD (2020) Autotriage—an open source edge computing raspberry pi-based clinical screening system. In: medRxiv. https://doi.org/10.1101/2020.04.09.20059840
20. Priyadarshini R, Barik R, Dubey H (2018) Deepfog: fog computing-based deep neural architecture for prediction of stress types, diabetes and hypertension attacks. Computation 6(4):62
21. Senthil Kumar A, Suresh G, Lekashri S, Ganesh Babu L, Manikandan R (2021) Smart agriculture system with E-carbage using IoT. Int J Mod Agric 10(1)
22. IoT in agriculture: 8 technology use cases for smart farming—S4IoT (2021) In: S4IoT. https://s4iot.com/iot-in-agriculture-8-technology-use-cases-for-smart-farming/. Accessed 14 Aug 2021
23. Cities (2021) In: United Nations sustainable development. https://www.un.org/sustainabledevelopment/cities/. Accessed 14 Aug 2021
24. IoT for smart cities: use cases and implementation strategies (2021) In: Scnsoft.com. https://www.scnsoft.com/blog/iot-for-smart-city-use-cases-approaches-outcomes. Accessed 14 Aug 2021

25. SMART CITIES (2021) In: Gudcltd.com. http://www.gudcltd.com/smart-cities. Accessed 14 Aug 2021
26. Smart infrastructures: essential for smart cities (2021) In: Nexus Integra (EN). https://nexusintegra.io/smart-infrastructures-essential-for-smart-cities/. Accessed 14 Aug 2021
27. Alfatihi S, Chihab S, Salih Alj Y (2013) Intelligent parking system for car parking guidance and damage notification. In: 2013 4th international conference on intelligent systems, modelling and simulation. https://doi.org/10.1109/isms.2013.35
28. Rahman M, Manik M, Islam M et al (2020) An automated system to limit COVID-19 using facial mask detection in smart city network. In: IEEE international IOT electronics and mechatronics conference (IEMTRONICS). https://doi.org/10.1109/iemtronics51293.2020.9216386
29. Jeong Y-S, Park JH (2019) IoT and smart city technology: challenges, opportunities, and solutions. J Inf Process Syst 15(2):233–238
30. Kubendiran M, Singh S, Sangaiah AK (2019) Enhanced security framework for e-health systems using blockchain. J Inf Process Syst 15(2):239–250. https://doi.org/10.3745/JIPS.04.0106
31. Choi J, Ahn S (2019) Scalable service placement in the fog computing environment for the IoT-based smart city. J Inf Process Syst 15(2):440–448. https://doi.org/10.3745/JIPS.03.0113
32. Park DM, Kim SK, Seo YS (2019) S-mote: SMART home framework for common household appliances in IoT network. J Inf Process Syst 15(2):449–456. https://doi.org/10.3745/JIPS.03.0114
33. Halder T, Karforma S, Mandal R (2019) A block-based adaptive data hiding approach using pixel value difference and LSB substitution to secure e-governance documents. J Inf Process Syst 15(2):261–270. https://doi.org/10.3745/JIPS.03.0111
34. Ding MJ, Zhang SZ, Zhong HD, Wu YH, Zhang LB (2019) A prediction model of the sum of container based on combined BP neural network and SVM. J Inf Process Syst 15(2):305–319. https://doi.org/10.3745/JIPS.04.0107
35. 7 industrial IoT applications (2021) In: Nexus Integra (EN). https://nexusintegra.io/7-industrial-iot-applications/. Accessed 14 Aug 2021
36. Wang Q, Zhu X, Ni Y et al (2020) Blockchain for the IoT and industrial IoT: a review. Internet Things 10:100081. https://doi.org/10.1016/j.iot.2019.100081
37. Liji Thomas M (2021) Blockchain applications in healthcare. In: News-Medical.net. https://www.news-medical.net/health/Blockchain-Applications-in-Healthcare.aspx. Accessed 14 Aug 2021
38. Blockchain in supply chain management | ConsenSys (2021) In: ConsenSys. https://consensys.net/blockchain-use-cases/supply-chain-management/. Accessed 14 Aug 2021
39. Blockchain in the energy sector: uses and applications | ConsenSys (2021) In: ConsenSys. https://consensys.net/blockchain-use-cases/energy-and-sustainability/. Accessed 14 Aug 2021
40. 8 blockchain startups disrupting the agricultural industry (2021) In: StartUs Insights. https://www.startus-insights.com/innovators-guide/8-blockchain-startups-disrupting-the-agricultural-industry/. Accessed 14 Aug 2021
41. https://manufacturingglobal.com/smart-manufacturing/role-blockchain-manufacturing (2021) Accessed 14 Aug 2021
42. https://www.pwc.com/us/en/industries/industrial-products/library/blockchain-industrial-manufacturing.html (2021) Accessed 14 Aug 2021
43. Gray C (2021) Energy consumption of internet of things applications and services. In: Minerva-access.unimelb.edu.au. https://minerva-access.unimelb.edu.au/handle/11343/224197. Accessed 14 Aug 2021
44. Digicatapult.org.uk (2021) https://www.digicatapult.org.uk/news-and-insights/blog/iot-interoperability-challenges-and-opportunities. Accessed 14 Aug 2021
45. IoT security issues in 2021: a business perspective (2021) In: Thales Group. https://www.thalesgroup.com/en/markets/digital-identity-and-security/iot/magazine/internet-threats. Accessed 14 Aug 2021

46. Success C (2021) Why IoT product development is complicated—and what you need to succeed | Flex. In: Flex.com. https://flex.com/resources/why-iot-product-development-is-complicated. Accessed 14 Aug 2021

Cumulative Neural Network Classification and Recognition Technique for Detecting and Tracking of Small-Size Garbage Materials

Vishal Pattanashetty, Suneeta Bhudhihal, K. Shamshuddin, Shweta Kore, and Shraddha Hiremath

Abstract A review of current advanced classification and recognition techniques, as well as recent technological advances in machine vision, shows that the application of a cumulative neural network (the most advanced machine learning algorithm in) for the waste management is an area of research that remains unexplored. The neural network used to prove the notion is referred to as a proof-of-concept neural network. A low-cost method for identifying and classifying recyclables, increase sorting efficiency, lower human workload, and boost to better comprehend the waste data network revolution how complexes of neural networks may change the industry of waste management. Using only color images of input waste, the system was able to classify objects with an accuracy of up to 90° by type of material (paper, glass, cardboard, metal, and plastic). The potential implementation of the recycling algorithm was assessed in terms of economic, social, commercial, and environmental performance, under the concept of integrated and sustainable waste management. When CNN-based systems are compared with existing waste management technologies, it has been found that they have the potential to modify extensive, semi-reversible manufacturer liability programs, and can change. The economy underpins all recycling.

Keywords Garbage detection · Machine learning · Image processing · Waste · Waste sorting · Analyzing waste · Segregation of waste · CNN · Machine learning

V. Pattanashetty (✉) · S. Bhudhihal · K. Shamshuddin · S. Hiremath
SoECE, KLE Technological University, Hubli, Karnataka, India
e-mail: vishalbp@kletech.ac.in

S. Bhudhihal
e-mail: suneeta_vb@kletech.ac.in

K. Shamshuddin
e-mail: shamshuddin@kletech.ac.in

S. Hiremath
e-mail: shraddha_h@kletech.ac.in

S. Kore
JSS College Dharwad, Dharwad, Karnataka, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_14

1 Introduction

Garbage management is a big problem in India that impacts a huge number of people in a variety of ways. With over 1.50 lakh metric ton (MT) of solid waste produced every day, India is buried behind a mountain of garbage. Worse, garbage accounts for almost 90% of the total amount (135,000 MT each day). Almost, 15,000 MT of rubbish is exposed every day, resulting in roughly 55 lakh MT of solid waste thrown in open locations each year, resulting in "serious" pollution. Only, 20% of all garbage collected is processed (27,000 MT per day), with the remaining 80% (108,000 MT per day) sent to landfills [1].

Researchers have offered litter reduction strategies in a number of ways. Despite this, the importance of actual behavioral intervention in litter reduction has received little research attention [2]. Based on a review of 50 studies, the five major approaches (environmental design, prompts, clean up the previous litter or beautification, environmental education, and environmental engagement) were found to have an influence on a person's littering habit [3].

2 Related Work

The study looks at existing state-of-the-art identification and sorting approaches as well as recent advancements in machine vision technology. The research includes traditional recycling systems so that the machine vision techniques could be compared to them [4]. While the study's goal was to cover techniques that could be applied to a variety of waste streams, it acknowledged that there was a tendency toward municipal solid waste owing to the greater number of research articles published in this sector. However, because solid waste is by far the most complicated and diversified waste stream, any waste development that makes categorization easier would necessarily apply to waste streams. Industrial trash, electronic waste, construction waste, and demolition debris are examples of other waste.

2.1 Manual Sorting

Manual labor is more cost-effective in underdeveloped nations than automated methods. The aggregate recovery efficiency of manual sorting is 91%. Considerably, more than any known automated approach, due to humans' unique ability to recognize materials in a relatively short amount of time. Manual sorting is usually limited to a last quality control step in industrialized nations. Because of budgetary restrictions, human weariness, and exposure to dangerous chemicals, it is typically avoided if feasible [5].

2.2 Machine Vision System

All machine vision systems, as illustrated in the graphic, image acquisition is handled by sensors, while image processing is handled by algorithms. For image processing, a device that removes the things that the vision system identified and found is added to the system for trash sorting [6].

2.3 Aerial Survey

Aerial surveys were done in Cambodia in October 2019 utilizing a DJI 4 Phantom Pro photography UAS with a 20 MP RGB (red, green, and blue color scheme) picture sensor over PhnomPenh, Sihanoukville, and Siem Reap (Fig. 1). Plastic trash is seen drifting, trapped in plants, washed ashore on beaches, and gathered to form plastic river carpets in Fig. 1b–d. Images were captured at a resolution of 4864 × 3648 pixels, with ISO settings ranging from 100 to 400 pixels, and shutter speed and aperture set to automatic. The image sensor's nadir viewing angle was 0 at a flying height of 6 m and a vertical GPS hover accuracy of 0.5 m.

The flight height was determined after analyzing imagery from pre-flight testing at altitudes ranging from 3 to 60 m. It provided adequate wide area coverage with sufficient object resolution (length 2.5 cm). A Topcon GR-5 global navigation satellite receiver system was used to identify the ground control locations. The points were used to improve image mosaicking and geo-referencing. The collected aerial photos were post-processed with Pix4Dmapper version 4.5.3. The project included automated point cloud densification, 3D mesh generation, digital surface modeling, orthomosaic, and digital terrain modeling. In certain pictures, a visual meter scale was used to supplement size estimations in photographed settings. The photos were not subjected to any atmospheric adjustment.

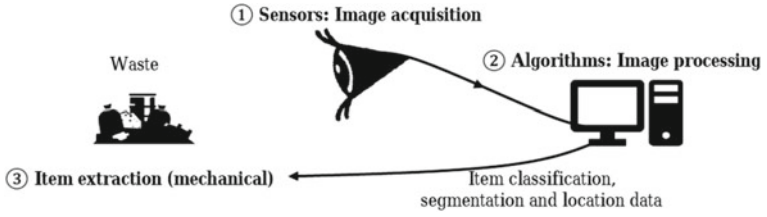


Fig. 1 Typical machine vision architecture

2.4 Detection and Quantification of Plastic Litter Algorithms

True color RGB photos were split into tiles of 100 × 100 × 3 pixels and 50 × 50 × 3 pixels throughout the theatrical survey. The tile size was determined based on the findings of a prior study (Martin et al. 2018). The plastic litter detector (PLD-CNN) technique is used to analyze the 100 × 100 × 3 pixel tiles. It was taught how to distinguish between (i) water, (ii) vegetation, (iii) litter-low, (iv) litter-high, (v) 11sand, and (vi) other tile targets (Fig. 2). Tiles with less than three objects were labeled litter-low, and those with three or more were labeled litter-high. After seeing that litter objects in images ranged from almost none in rivers with natural surroundings to hundreds of litter objects in plastic river carpets, we devised these parameters to ensure that the algorithm could count as many litter items as possible. From the 100 × 100 × 3 pixel tiles used by PLD-CNN, a division of 50 × 50 × 3 pixel tiles was chosen for the plastic litter quantifier method. The purpose of using a divider was to enhance the algorithm and therefore decrease information loss. PLQ-CNN only evaluated tiles with any amount of garbage. PLQ-CNN was also taught to recognize and count individual litter bits inside these tiles. In addition to the six PLD-CNN categories, PLQ-CNN created 18 classes, including cans, cartons, plastic bags, bottles, cups, canisters, polystyrene packaging, shoes, Styro-foam, strings, and fabrics (Fig. 2). Plastic bags were used to package sweet wrappers, noodle packages, and chips bags, which were divided into large and small bags. The items were recognized using the new international marine litter classification technique developed by the US National Oceanic and Atmospheric Administration. Following the theatrical survey, a visual assessment of these items or plastic objects was conducted as part of the cleanup procedure [7].

Also, laser-induced breakdown spectroscopy (LIBS), optical based sensors, X-ray sensors, ultrasound and noise and electromagnetic sensor (EMS) are used for detection.

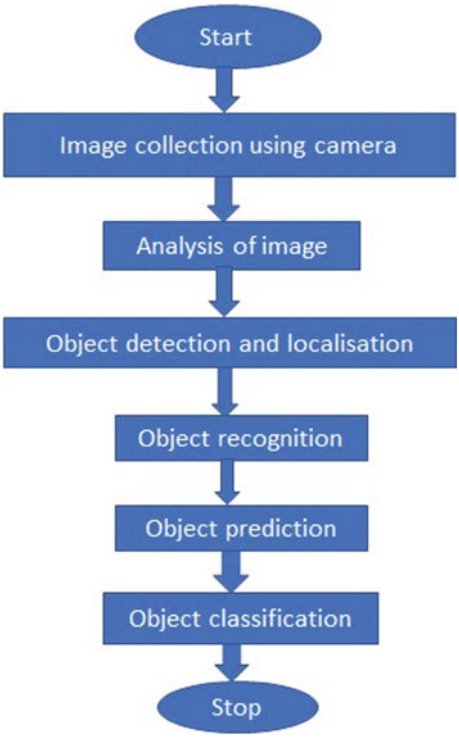
Fig. 2 Aerial view of garbage

3 Proposed Method

3.1 Object Detection

We utilized the TensorFlow API to build the algorithm, as well as Jupyter Notebook to interact with it and an Amazon Web services Ubuntu server to train it. Using TensorFlow's pretrained model and the COCO dataset, the team was able to lay the groundwork for the project's object detection component. As seen in the above image, Prototype 1's resultant technique boxed almost every item it identified inside an input image. Each recognized object is surrounded by a box. The confidence indication next to each identified object is a percentage that indicates how accurate that detection was. For example, the algorithm is 76% positive that it recognized a kite in the image provided. A grading system is also in existence, albeit it is not standardized. The number of items spotted in each image in Prototype 1 was scored from one to four. This functionality paved the way for the eventual inclusion of image ranking based on litter severity in Prototype 2. For the initial prototype, the image's final output boxed the elements with the accuracy percent and a rating (Fig. 3).

Fig. 3 Flowchart



3.2 Dataset

We realized that the algorithm's litter identification skills required a large dataset dedicated to litter after hours of research. Pictures from Google Streets were used to build the dataset that the team eventually came up with. Pulling a string with your hands, we created a script to automate the Google Streets image extraction procedure. Hundreds of thousands of photographs would be collected automatically. The crew needed to identify the pictures after they were acquired in order to start the training process. As a consequence, 10,000 pictures from Google Streets were used to build the dataset. Keep Earth Beautiful organized a photo identification contest for a group of children. This group utilized the labeling software that we provided. The students initially classified the photographs into two categories: those with litter and those without litter. They then identified all of the pictures using garbage. Once the students had done labeling, we needed to double-check the labels for accuracy. During the verification process, the team noticed that some of the recognized images were incorrectly labeled, rendering them worthless. Approximately, 3000 images were considered appropriate for the dataset at the conclusion of the operation. Although not all of the photographs were recognized, the 3000 that was proved to be extremely useful, given the total of 9000 items of garbage.

3.3 Training

The goal of Prototype 2 was to finish the garbage recognition part of the algorithm. However, the team would have to train the labeled dataset to match their criteria before deploying the recognition. The researchers split the dataset in half, utilizing 80% of the images for training and the remaining 20% for testing. After training the algorithm to a certain number of "steps," we used a checkpoint file to evaluate the program's litter detection accuracy. These checkpoint files informed the team when it was time to stop training and evaluate the algorithm's progress. If the team believes the algorithm has been over trained, the previously stored checkpoint file will be used (Figs. 4, 5, and 6).

3.4 Litter Detection

After the algorithm had been trained, we put it to the test using a variety of images. We also utilized images from the collection that did not contain garbage. The organization is in desperate need of help. To guarantee that the software does not box them even if there are no other things in the image. Litter may be seen in the image. After hours of testing and training, the crew was finally satisfied with the result. Only, box pieces

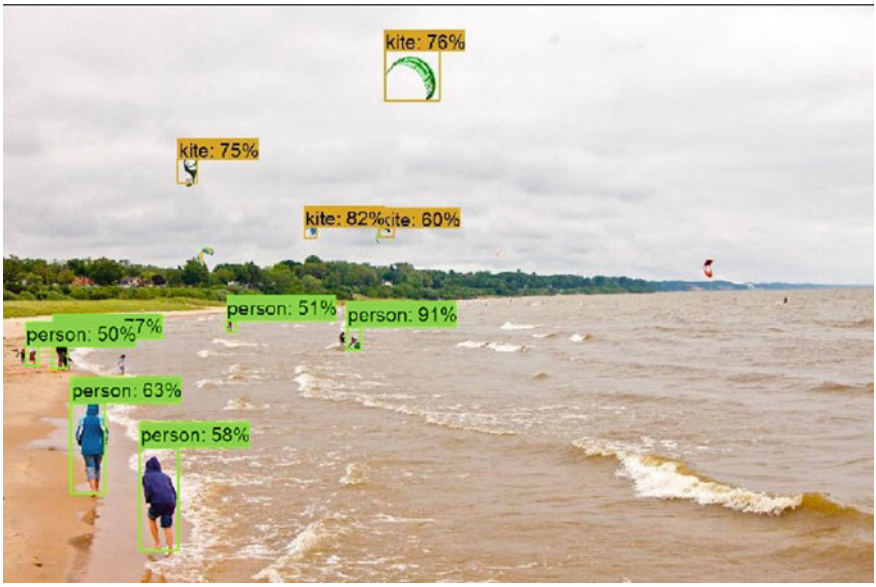


Fig. 4 Sample 1

Fig. 5 Sample 2

of litter with a height of over a meter appear in the algorithm's final output. There is a 50% confidence level for the image, as well as a rating (Fig. 7).

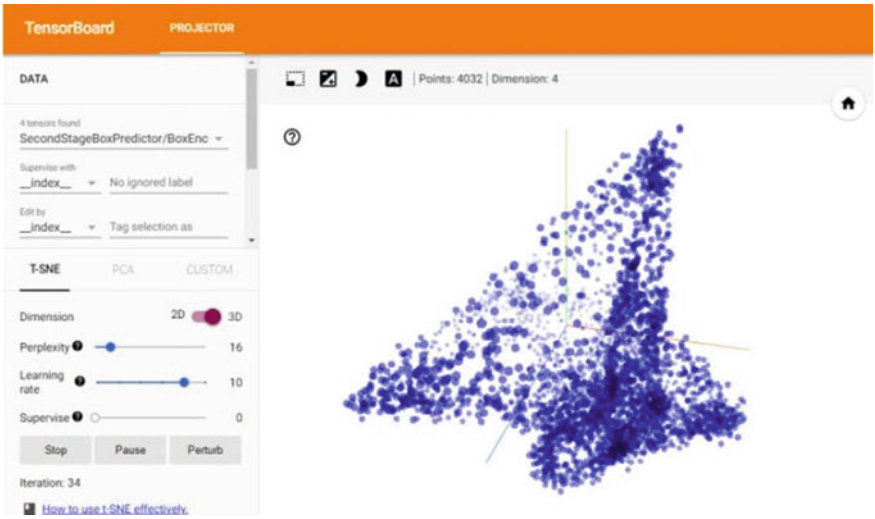


Fig. 6 Training dataset

Fig. 7 Detection

4 Results

We are able to detect various sizes of garbage materials varying from 0.5-cm to 20-cm feet (Figs. 8 and 9).

Fig. 8 Garbage detection result 01

Fig. 9 Garbage detection result 02

5 Conclusion and Future Scope

5.1 Conclusion

Most regions have a significant demand for cleanliness when it comes to trash detection. This new innovative method of the trash-detecting module paired with iterative approaches would assist the vast majority of establishments that currently rely on traditional cleaning processes. The results of this research will be utilized to establish priorities and ensure that all places are kept as clean as possible.

5.2 Future Scope

More study should be conducted, this time concentrating on other types of enterprises and sectors. To put it another way, more study in a range of empirical circumstances will be needed to strengthen the validity of the concept.

References

1. Qing L, Yao C (2003) Real-time concepts for embedded systems; Fa-deng G, Min-xian H (2010) Study on the detection of locomotive driver fatigue based on image. In: 2010 2nd international conference on computer engineering and technology. IEEE
2. Pattanashetty VB, Mane V, Iyer NC, Kore S (2022) Traffic rules violation detection system. In: Joshi A, Mahmud M, Ragel RG, Thakur NV (eds) Information and communication technology for competitive strategies (ICTCS 2020). Lecture notes in networks and systems, vol 191. Springer, Singapore. https://doi.org/10.1007/978-981-16-0739-4_8
3. Pattanashetty VB, Mane V, Hurkadli SS, Iyer NC, Kore S (2022) Lane detection for visual assistance of self-driving vehicles for structured and unstructured roads. In: Joshi A, Mahmud M, Ragel RG, Thakur NV (eds) Information and communication technology for competitive strategies (ICTCS 2020). Lecture notes in networks and systems, vol 191. Springer, Singapore. https://doi.org/10.1007/978-981-16-0739-4_27
4. Pattanashetty VB, Iyer NC, Viswanath HL, Kore S (2019) Inclusive device to connect people with the surrounding world. In: Fong S, Akashe S, Mahalle P (eds) Information and communication technology for competitive strategies. Lecture notes in networks and systems, vol 40. Springer, Singapore. https://doi.org/10.1007/978-981-13-0586-3_49
5. Pattanashetty VB, Iyer NC, Viswanath HL (2018) Intellectual conveyance structure for travellers. In: Saeed K, Chaki N, Pati B, Bakshi S, Mohapatra D (eds) Progress in advanced computing and intelligent engineering. Advances in intelligent systems and computing, vol 564. Springer, Singapore. https://doi.org/10.1007/978-981-10-6875-1_37
6. DevKit8000 Evaluation Kit manual 51. http://dlib.net/facelandmarkdetectionex.cpp.htm
7. Pattanashetty VB, Iyer NC, Ibal P, Bagewadi A (2015) Smart driving assistance. IEEE Int Transp Electrification Conf (ITEC) 2015:1–4. https://doi.org/10.1109/ITEC-India.2015.7386870

Performance Analysis of OFDM-IM Using DLD

Chilupuri Anusha and S. Anuradha

Abstract The proposed work presents an advanced signal detection methods which is a wide challenge to model it for OFDM-IM. However, the complexity at the receiver side is widely increased due to introduction of Index Modulation (IM), and for IM, it also needs to know the Channel State Information (CSI), which increases the complexity and it also results to system overhead. The deep learning-based detector (DLD) improves the system performance and avoids the system overhead, comparing to the tradition detectors like Maximum Likelihood (ML), Greedy Detector (GD), Log Likelihood Ratio (LLR), etc. The proposed deep learning-based detector (DLD) uses deep neural network (DNN) with fully automated connecting layers is used for detecting the data bits at the receiver of OFDM-IM system. Firstly, the deep learning detector (DLD) is trained offline by collecting the data sets of the simulated results for improving the BER performance and accordingly the model gets trained to use it for online detection of OFDM-IM signal at the receiver. The results proves that the deep learning-based detector provides an adequate BER performance with a minimum runtime than the traditional detecting methods.

Keywords OFDM-IM · Channel state estimation · Deep neural network · Deep learning · Performance of BER

1 Introduction

OFDM-IM [1] has been appeared as an ensuring multi-carrier system destined to substitute traditional OFDM system. In OFDM-IM, along with the active subcarriers their indices are used to send the information bits as there is no need of using any additional power or bandwidth, this makes the system more efficient and reliable than

C. Anusha (✉) · S. Anuradha
National Institute of Technology Warangal, Warangal, Telangana, India
e-mail: ca721059@student.nitw.ac.in

S. Anuradha
e-mail: anuradha@nitw.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_15

the OFDM. Apart from this, OFDM-IM also invents a new method for efficient use of spectrum and to make a system more reliable while selecting the active subcarriers. Looking for their results, OFDM-IM has a great part of research attentiveness and it is also noted as a most engaged system in research now a days. In paper [2, 3] has shown the methods for analyzing the performance of bit error rate in an imperfect channel state information. Different methods are introduced to progress the performance of OFDM-IM system, shown in [4] and the other references. BER performance is analyzed by using the traditional greedy detector was shown in [5] and the simulated results are in [6].

The traditional detection methods of OFDM-IM results better performance but also results in high complexity comparing to OFDM system and it is difficult to overcome this by using the traditional methods by sustaining the optimal performance of the system. Maximum Likelihood detector results in optimal BER performance with rise in complexity, to overcome this a low complexity LLR ratio was introduced in [1] to attain an optimal near ML performance but it needs a specific understanding on power spectral density of received noise. The greedy detection method also involves high complexity than OFDM detector and it also suffers more amount of performance loss than the ML detector. However, in order to avoid and deal with these downsides of the traditional detection methods, a DLD is used.

Over the last few years, deep learning [7] has become a fuzz word in the technical community. At all times we usually hear about deep learning in news and applications of deep learning in various fields, like natural language processing and computer vision. Now a days it has also applied in the field of communication systems to avoid physical layer problems. Deep learning algorithms uses a deep neural network which tells the relation between the inputs and outputs. In communications systems they proposed a deep learning auto encoder [8] to replace both the transmitter and the receiver units with the deep neural networks. Deep learning also applied in OFDM systems to reduce the PAPR [9] and also used to design a channel estimation system and the signal detector [10] and this auto encoder system [11] also inspected in OFDM systems in existence of hardware impairments.

This work proposed a deep learning detector (DLD) for OFDM-IM to result in better performance and to reduce the complexity over the prevailing detection methods. In deep learning-based detector, before providing the data to the DNN the information about the channel and the received signal are processed initially based on the basic understanding of OFDM-IM. The projected DNN structure requires two fully connected layers to execute the recognition of signal over Rayleigh fading channels. Furthermore, to walkout the trade-off among the complexity and the channel performance, number of nodes can be accustomed accordingly. Firstly, using the simulated data, deep learning-based detector will be trained offline in order to minimize the BER and then it is used as available online detector with extremely less execution time. The results revealed that the deep learning-based detector gives a proper BER for both the perfect and the imperfect CSI with a less runtime than the existing detector methods.

The outline of the script is assembled as follows. Model of the system is explained in Sect. 2, Projected DLD explained in Sect. 3.1, outcome of the simulation is displayed in Sect. 3.3 and the inference in Sect. 4.

2 System Model

Among various multi-carrier modulation techniques, OFDM-IM is very optimistic technique, where all the OFDM subcarriers are triggered to transmit signal constellations and also carries an additional information to maintain high performance and spectral efficiency. In this document we are considering an OFDM-IM system with N_T subcarriers, each of N subcarriers are grouped into L groups, i.e., $N_T = NL$. The processing of signal for every group is same and self-determining at the transmitter side. To make it simple here, we are considering only one group, while transmitting the group only a few out of 'N' subcarriers, say 'K' are activated to send all the input data bits, say p, of M array modulation size that includes symbols which carries data and indices of active subcarriers, are given by '$p1 = K \log_2 M$' and '$p2 = \left\lfloor \log_2 C(N, K) \right\rfloor$', i.e., $p = p1 + p2$. Mapping of combination of active subcarriers to the data bits is implemented by lookup table [1] or by combinational method. The transmitted vector is constructed by allotting the data symbols which are non-zero, to active subcarriers. Let us say the transmitter vector is $x = [x_1, \ldots, x_N]$, x_i is not equal to zero only if the subcarrier 'i' is active, else ways zero. Bit to symbol mapping is denoted by a function $x = f_{\text{OFDM-IM}}(b)$, '$b$' denotes the series of incoming bits in one group.

The received output signal is given by

$$Y(n) = h(n) \odot x(n) + w(n) \tag{1}$$

Here, $\odot$ stands for circular convolution, $h(n)$ is a channel in this paper, considering it as Rayleigh fading channel $h = [h_1, \ldots, h_N]$, $w(n)$ denotes the noise considering AWGN noise with its entries ni $\text{CN}(0, \sigma^2)$, $i = 1, \ldots, N$. The average SNR for the received output signal is given by $\gamma^- = \text{Es}/\sigma^2$, where '$\sigma^2$' is the mean square error of the received signal and 'Es' is the mean of transmitted M array symbol energy.

As the signal detection methods of OFDM-IM like ML, GD and LLR is having downsides in terms of system performance and complexity. To avoid these negativities a new detector is going to implement based on deep learning, for modeling an efficient DNN for estimating the channel and detection of signal it requires two stages. The neural network is trained offline by the simulated results which are already obtained under various channel conditions with different information sequences. In the online deployment step, the DNN model will generate the output which recovers the transmitted data without estimating the wireless channel.

3 Deep Neural Network

In this segment, we first understand basic model of proposed DLD and then the model training and online deployment of the trained model.

3.1 Basic Model of DLD for OFDM-IM

Deep learning methods has been used in different range of application like robotics, CV, ML, natural language processing, speech processing and, etc., to improve the system performance. The model of deep neural network [12] is presented in Fig. 1. In general, DNN's are the special versions of ANN with increase in hidden layers in order to improve the identification skill. All the hidden layers of the network comprises multiple neurons, and has an output of irregular function of a weighted sum of neurons of its former layer, as shown in Fig. 1. This non-linear function may be a rectifier linear unit (Relu) function, or a hyperbolic tangent (Tanh) function or a sigmoid function.

The deep learning-based detector is similar to already present detection methods of OFDM-IM, i.e., the information of the channel is expected to be recognized at the destination side. Hence, the output signal 'y' and channel 'h' are taken as the main inputs of deep neural network. These 'y' and 'h' are processed initially before

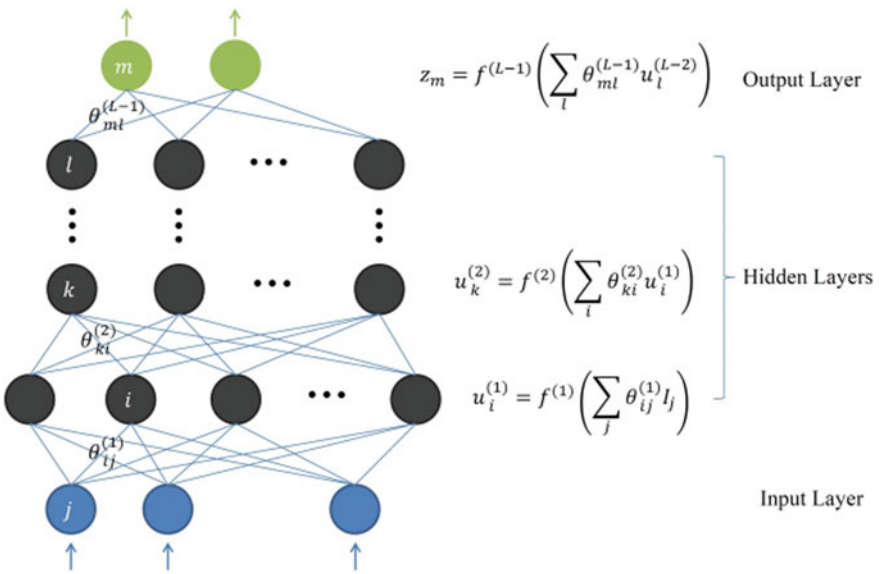


Fig. 1 A basic deep learning model

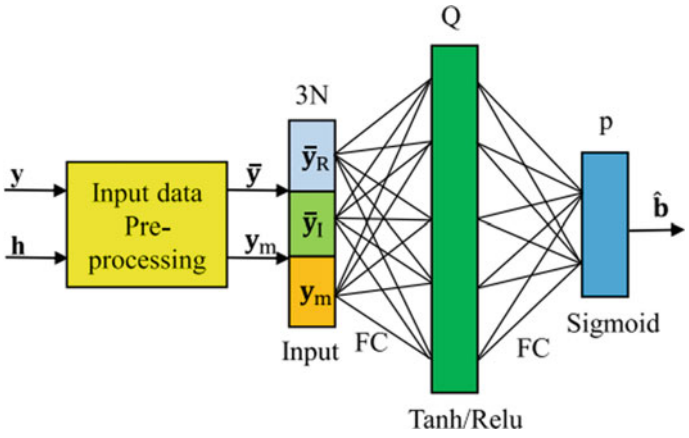


Fig. 2 Proposed DLD for OFDM-IM

entering into the DNN structure, to get an optimal detection performance as an outcome.

The structure of proposed DLD is shown in Fig. 2. ZF equalizer is applied to get an equalized output signal can be given as $y^1 = y \odot h^{-1}$, it improves the reconstruction on M-ary Modulated symbols and the energy of the output signal given by $y_m = \left[|y_1|^2, \ldots, |y_N|^2\right]$. y^1 and y_m are combined to give the input to neural network. In greedy detector, y_m is used for decoding the indices of the subcarrier based on the basic knowledge of greedy detector to improve the performance of signal detection. The output of DNN can be displayed as

$$\hat{b} = f_{\text{Sigmoid}}\left(W_2 f_{\text{Tan/Relu}}(W_{1Z} + b_1) + b_2\right) \tag{2}$$

Here, $\hat{b}$ is the estimation of transmitted data, W_1 and W_2 are the weights of fully connected layers and b_1 and b_2 are the basis of fully connected layers. In the proposed deep neural network, the length of inputs and outputs are determined by system parameters like N, K and M, hidden layer length need to be selected properly to get the better performance of system. The number of transmitted bits must be directly proportional to the length of hidden layer in order to give better performance. The merits of the projected DLD is the length of the input vector highly depends on N parameter and on K and M it is less dependent, this makes the system less complex comparing to other existing detectors as they depend on all the three parameters.

3.2 Training the System and Deploying Online

Deep learning-based detector is first trained by using the offline data which is obtained by simulation of OFDM-IM and various bits 'b' are generated randomly to get

the series of vectors transmitted as $x = f_{\text{OFDM-IM}}(b)$. Later all these vectors are transmitted to target receiver through the Rayleigh fading channel with AWGN noise. There is a provision of changing the channel and noise vectors according to the system. The output signal received and the channel vectors are processed before to get the input data set for the corresponding 'b' bits. Rather the number of training data samples improves, the accuracy of the system performance will also get improved and also prevents overfitting.

For training the system efficiently, SNR level must be selected properly as the system performance depends more on it. And the trained model is used online for detection of OFDM-IM. The details of SNR and the modulation selected are given in the simulation results.

3.3 *Simulation Outcomes*

We deliver the simulation results in terms of BER versus SNR Performance of proposed DNN under various channel states for $M = 4$.

Figure 3 evaluates the BER performance of the DLD under different channel states specially, perfect and imperfect CSI, when $(K, N, M) = (1, 4, 4)$ and we can observe that the performance is better under perfect CSI. We also compared the performance of DLD with the other detectors shown in Figs. 4 and 5.

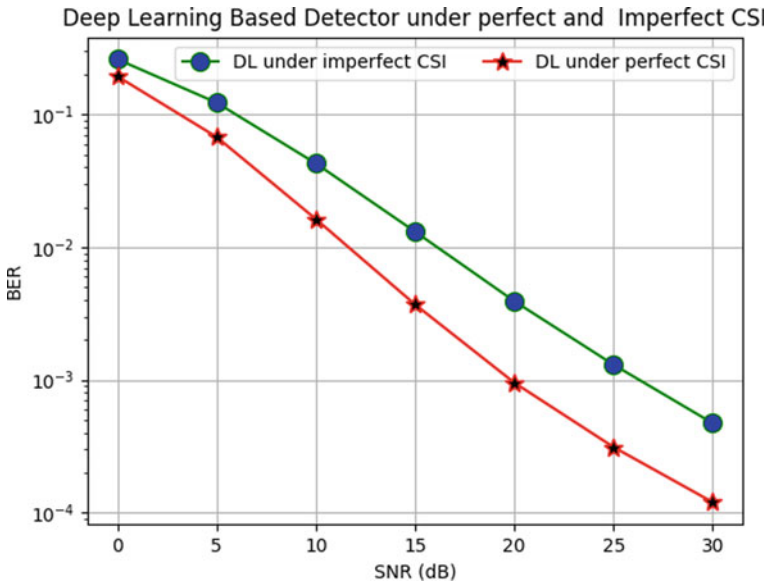


Fig. 3 Comparison of BER performance using DLD under various channel for $M = 4$

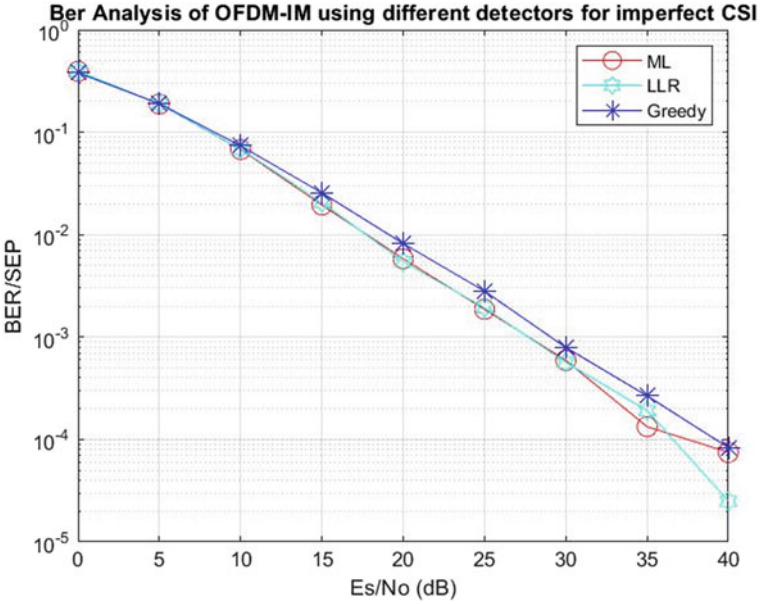


Fig. 4 BER performance using ML, LLR and greedy detector under imperfect CSI

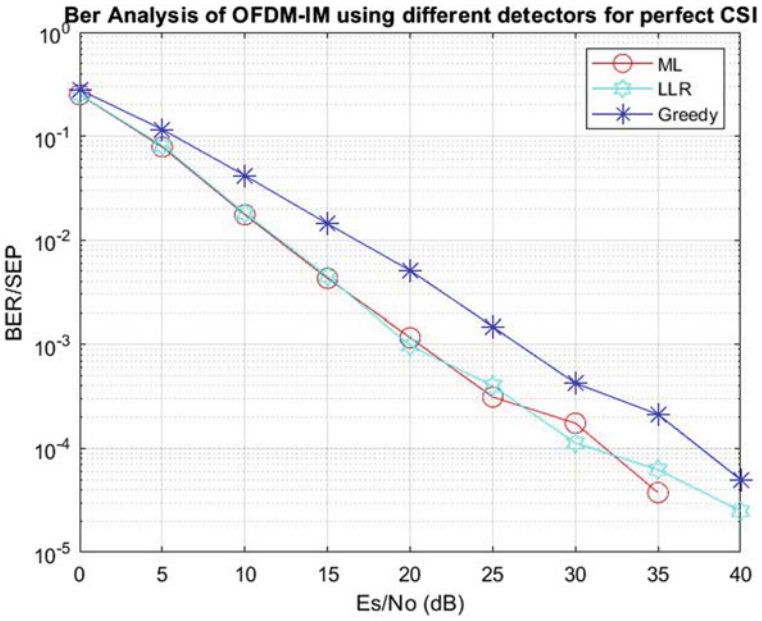


Fig. 5 BER performance using ML, LLR and greedy detector under perfect CSI

Table 1 Comparison of SNR versus BER for detection without deep learning (ML, LLR and GD) and with deep learning (DLD) under perfect and imperfect CSI

	BER under perfect CSI				BER under imperfect CSI			
	Without DL			With DL	Without DL			With DL
SNR (dB)	ML	LLR	GD	DLD	ML	LLR	GD	DLD
5	0.0795	0.075	0.116	0.0687674	0.19	0.186	0.188	0.1208599
15	0.0048	0.0042	0.01432	0.0038725	0.020	0.019	0.025	0.0129724
25	0.0004	0.00036	0.00167	0.0003275	0.008	0.0059	0.006	0.0012724

From Figs. 3, 4 and 5 we can notice that the working of DLD is better and optimal comparing with the other detectors without deep learning.

Table 1 shows the comparison among various detectors without deep learning and with deep learning. When we use detection methods without deep learning, in the performance evaluation we can observe that the LLR detector gives better performance than the other detectors that being the case consider for perfect channel state, at SNR = 15 dB, we got BER as 0.0048 for ML, 0.0042 for LLR and 0.01432 for GD. Among these detectors LLR has the better performance of BER. By comparing this LLR with DLD, consider at same SNR = 15 dB, BER = 0.00387 for DLD, i.e., even better than the LLR. Hence, we can infer that the performance of DLD is better than the detector without deep learning, accordingly we can look-at the performance of detection without deep learning and with DLD for imperfect channel state. From both the observations we can infer that the proposed DLD gives best performance than the other detectors.

Thereby we also applied DLD for various $M = 4, 8, 16$ under perfect and imperfect CSI and the simulated results are shown in Fig. 6.

4 Inference

In this Script, a novel detection method is evidenced for OFDM-IM. The DLD gives efficient performance by using deep learning in detection of OFDM-IM signal in terms of SNR versus BER performance. The projected DLD has pre-processed the received output signal and the channel vectors depending on the basic awareness of equalizing techniques, like ZF, MMSE and before training DNN with fully connected layers, energy of the signal has to be detected to recover the transmitted bits. It is shown that once the model is trained, the DLD can stand to use it in online with minimum runtime by achieving best performance. We trust that these significant merits of DLD will motivate in future work of relating deep learning to future wireless communications. In future work, the performance can be analyzed by applying different machine learning algorithms in place of DNN like complex CNN, etc.

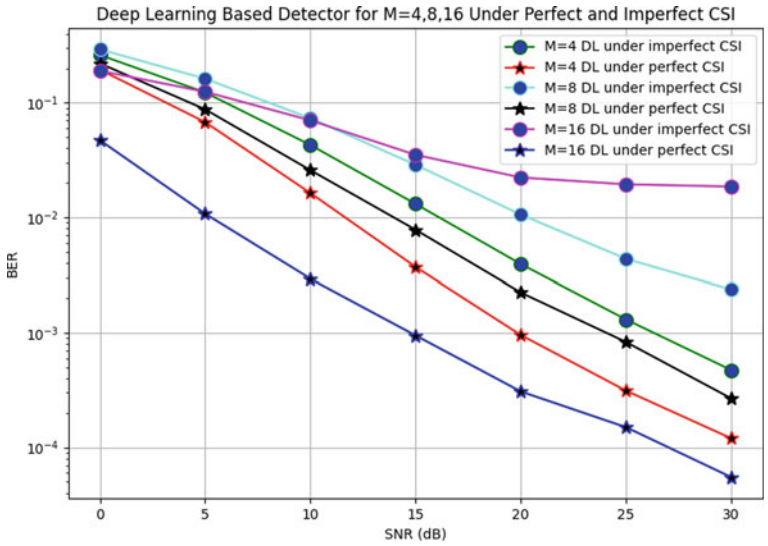


Fig. 6 Simulation results of deep learning detector (DLD) for $M = 4, 8, 16$ under perfect and imperfect CSI

References

1. Basar E, Aygölü U, Panayırcı E, Poor HV (2013) Orthogonal frequency division multiplexing with index modulation. IEEE Trans Signal Process 61(22):5536–5549
2. Luong TV, Ko Y (2018) Impact of CSI uncertainty on MCIK-OFDM: tight closed-form symbol error probability analysis. IEEE Trans Veh Technol 67(2):1272–1279
3. Luong TV, Ko Y, Choi J (2018) Repeated MCIK-OFDM with enhanced transmit diversity under CSI uncertainty. IEEE Trans Wireless Commun 17(6):4079–4088
4. Wen M, Ye B, Basar E, Li Q, Ji F (2017) Enhanced orthogonal frequency division multiplexing with index modulation. IEEE Trans Wireless Commun 16(7):4786–4801
5. Crawford J, Ko Y (2015) Low complexity greedy detection method with generalized multicarrier index keying OFDM. In: Proceedings of the IEEE PIMRC
6. Luong TV, Ko Y (2017) A tight bound on BER of MCIK-OFDM with greedy detection and imperfect CSI. IEEE Commun Lett 21(12):2594–2597
7. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117
8. O'Shea T, Hoydis J (2017) An introduction to deep learning for the physical layer. IEEE Trans Cogn Commun Netw 3(4):563–575
9. Kim M, Lee W, Cho D-H (2018) A novel PAPR reduction scheme for OFDM system based on deep learning. IEEE Commun Lett 22(3):510–513
10. Ye H, Li GY, Juang B-H (2018) Power of deep learning for channel estimation and signal detection in OFDM systems. IEEE Wireless Commun Lett 7(1):114–117
11. Felix A, Cammerer S, Dörner S, Hoydis J, Brink ST (2018) OFDM autoencoder for end-to-end learning of communications systems. In: Proceedings of the IEEE, SPAWC 2018
12. Van Luong T, Ko Y, Nguyen DHN, Matthaiou M (2019) Deep learning-based detector for OFDM-IM. IEEE Wireless Commun Lett 8(4):1159–1162

Blockchain Technology—A Brief Survey

F. Monica Rexy and G. Priya

Abstract Today's world is all about innovations. At present, any human cannot lead a normal lifestyle without the application of software technologies in their day-to-day activities. In such case, security has become a great threat to maintain sensitive records of people. Blockchain technology provides a solution in this case. A blockchain is indeed a public ledger that is open to all but without any control by a central authority. It is a technology that enables individuals and businesses to work together with trust and accountability. Cryptographic currencies like Bitcoin are one of the best-known blockchain applications. Blockchain technology is considered as the driving factor behind the next fundamental IT revolution. Several blockchain technology implementations are commonly accessible today, each with its particular strength for a particular application domain. This paper gives a brief introduction about various applications of blockchain technology in various domains.

Keywords Blockchain technology · Cyber risks · Information security

1 Introduction to Blockchain Technology

This twenty-first century is filled with innovations and technologies. Throughout our daily lives, with the growing need for technology, people are making themselves ready to support upcoming innovations. Existing innovative technology has cleared a path to our typical lifestyle, from using a remote in controlling electronic devices to using voice notes to indicate directions. Discoveries such as Bitcoin and Internet of things have gained momentum in the last decade, and there is yet another addition to the team, that is, blockchain technology [1]. The term blockchain was initially coined by Satoshi Nakamoto in 1992, but this technology actually came into picture

F. Monica Rexy (✉) · G. Priya
Vellore Institute of Technology, Vellore, Tamil Nadu 632014, India
e-mail: monicarexy.f@vit.ac.in

G. Priya
e-mail: gpriya@vit.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_16

only in 2008. In today's world, blockchain technology is emerging as recent trend to provide maximum security, transparency, and immutability in any transaction-based scenario.

Blockchain can simply be defined as a series of record that grows on continuously whenever a new transaction is added. This addition of a new block always happens on the rightmost end of a blockchain. The data once entered inside a block can never be altered or updater at any cost, thus it fulfills the property of immutability [2]. The structure of a blockchain is build using Merkle tree approach and it adds consensus to it. Consensus is a set of protocols that verifies how the transactional data are added and how it is then authenticated. The major advantage of Merkle tree in blockchain is that it provides greater hashing technique to provide security to the data. Bitcoin used blockchain technology to record all the transactions occurred in it. Blockchain is not only restricted to financial framework but also it is an excellent response for any framework that involves trust, i.e., keyless automobile system. In addition to that, Samsung and IBM recently opted for proof of concept which uses blockchain technology as its key technology in IoT. Blockchain also improves the quality of the systems that are used already in the industry [3].

1.1 Characteristics of Blockchain Technology

This blockchain technology has some special characters which promotes this technology over many other, and those special characters are explained as follows [4].

Peer-to-peer network: A peer-to-peer network is an infrastructure for decentralized networks that comprises of several computers that are connected together without a centralized database. Privacy is limited to only limited nodes in hierarchical networks. If the main server is targeted, the overall network security will be breached. On that contrary, this peer-to-peer network is considered to be better in such a case where no single point of failure occurs.

Distributed Ledger: This ledger is a computing system that records all the transactional details with respect to the sender and the receiver of a process. The distributed ledger is an information system that is shared through various systems within the network. Distributed ledger technology (DLT) describes about spreading the data to every user in a specific network. DLT comprises of three elements–data model (current ledger's status), transaction language (used to change the state of ledger), and protocol (used in building consensus). Blockchain is considered as one of the best examples of DLT. The exchange of information among all the clients enhances transparency and prevents any wrongdoing.

Consensus: Consensus is a mechanism that ensures every different node in the network comes to a conclusion about the present state of blockchain. Blockchain uses various algorithms and architectures to achieve consensus in different scenarios.

For an instance, Bitcoin applies proof of work concept in achieving consensus, and on the other hand, Ethereum started using proof of stake concept instead of proof of work algorithm.

Smart Contracts: Smart contract comprises of some set of protocols that is to be followed before and during the transactions in the blockchain network. Forget about smart contracts and blockchain technology for a second and assume a generalized contract. They contain certain conditions that must be met to make certain possible transactions (e.g., financial transactions). Let us take an example of selling a laptop. Here, a contract between the seller and buyer will be set like, the buyer will transact the money to seller only when the laptop works perfectly. In similar way, smart contracts also talk about the terms and conditions formatted between the send and receiver of nodes in the network.

1.2 Representation of a Block

Each block in a blockchain consists of three parts. They are data, hash, and hash of the previous block [5].

Data: The data part comprises of actual data or transactional record that is stored in the particular block.

Hash: It is the cryptographic hash function which is used to secure the data stored in that specific block. This hash value acts as an identification term that is derived after encrypting the data of that block. There are many cryptographic algorithms to perform this hashing. Some of them are SHA 1, SHA 256, MD 5, Keccak 256.

Hash of the Previous Block: The third part of data in a block contains the hash value of the previous block to maintain the continuity in the blockchain [12]. Because of this previous hash value, any block in the blockchain can be easily located and accessed. The pictorial representation of a single block in a blockchain is shown in Fig. 1.

Fig. 1 Representation of a single block [6]

1.3 Features of Blockchain Technology

Blockchain technology can be widely applied in many fields and it is mainly because of its unique features that no other technologies offer. The key features of blockchain include usage of ledger, immutability, chain of blocks, transaction validation, decentralization, and better security [7]. These features are briefly explained as follows.

Usage of Ledger: Blockchain technology uses a ledger to record all the transactions occurred in every block with its timestamp. Whenever a transaction is requested or completed, it is immediately recorded in this ledger. That is, ledger keeps track of all the actions taking place in the entire blockchain irrespective of any particular block. It is helpful in such a case when an admin or an end user attempts to checks or validate the actions took place in any block.

Immutability: The transactional data once entered into the block or into the ledger can never be altered or updated. If an attempt of alteration occurs, the hash value of the particular block and the previous hash value of the block changes which results in complete mess in the entire blockchain.

Chain of Blocks: This feature helps in maintaining the continuity and series of record in the blockchain network. This is managed by hash value and the previous hash value of blocks. It is helpful in minimizing the time taken to find out a particular block among several blocks in the blockchain network.

Transaction Validation: Whenever a new transaction is created, a copy of that transaction is sent to every node in the network. All the participating nodes and the miners in the network validate the transaction with respect to the consensus protocol. If the majority of the nodes accept the transaction, a block is created and added to the blockchain, else the transaction gets destroyed.

Decentralization: Decentralization is completely different from centralization. Decentralization offers better security and reliability when compared to centralized databases. In blockchain, the power of administration is distributed among several sections within the network. The section head is chosen and takes care of every action occurs in that specific section. The data loss or any attempt of fraudulent activity in one section does not affect the data in any other section. And decentralization offers quick decision-making. Representation of a centralized and decentralized network is shown in Fig. 2.

Better Security: The cryptographic hash function provides at most security by forming very complex hash value that challenges every hacker to think twice before making an attempt of hacking. By any chance, if an attempt is made, 51% of the system get collapsed resulting in compromising the whole network. And if a network of system is destroyed, the data can be retrieved from other network of systems, since it follows decentralized database.

Fig. 2 Representation of distributed, centralized, and decentralized network

2 Types of Blockchain Technology

There are four primary types of blockchain [7]. They are public blockchain, private blockchain, hybrid blockchain, and consortium blockchain.

Public Blockchain: This is said to be permission less blockchain since it is open source. It allows anybody to be engaged in as user, and the transactional details are also available to everyone in the network. This public blockchain is fully decentralized and always hosted on public servers. It is anonymous and highly resilient to threats.

Private Blockchain: It is partially decentralized and is normally controlled by a highly trusted organization. In this private blockchain, it is necessary to get authorized in order to join the network. All the transactions in this private blockchain are kept private, that is, only authorized users can view the transactions. This type of blockchain is always hosted on private servers and is considered as highly scalable [8].

Hybrid Blockchain: This is the combination of both public and private blockchains. This type of blockchain is controlled by some of the dominant nodes in the same network. It is partially decentralized. And here also, authorization is required in order to join the network and view the transactions.

Consortium Blockchain: This has the same property as private blockchain. In addition to that, consortium blockchain is considered as a subcategory of private blockchain system. Unlike private blockchain, consortium blockchain is controlled by a group, rather than a single entity.

3 Working Model of Blockchain

As already discussed, blockchain is a series of record which are immutable and time stamped. The working process of blockchain comprises of five phases [9]. They are

- Phase.1: Requesting a transaction,
- Phase.2: Block creation,

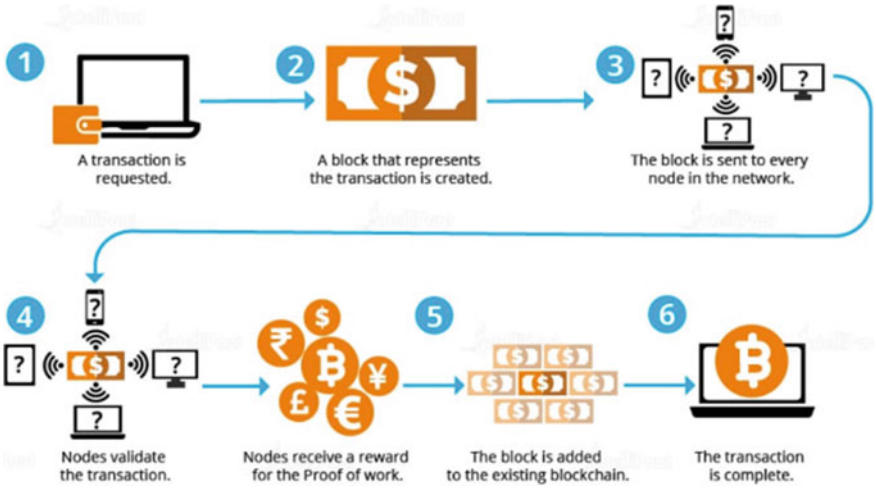


Fig. 3 Working process of blockchain technology [11]

- Phase.3: Sending copy of block,
- Phase.4: Block validation, and
- Phase.5: Adding block to blockchain.

The brief explanation of the process involved in each phase is explained below. And the pictorial representation of the working model of blockchain technology is shown in Fig. 3.

- Consider a scenario that the blockchain system has five blocks such as A, B, C, D, and E, respectively. If A needs to do a transaction to B, a transaction request is initiated and a block that represents the transaction is created [10].
- A copy of the requested block is sent to every participant (i.e., to B, C, D, and E) in the network.
- All the participating nodes validate the block with respect to the consensus protocol.
- If the majority of the nodes accept the transaction, the new block is added to the existing blockchain, else the block gets destroyed.
- Once the block is linked with the blockchain, the transaction is considered as complete. And by this way, a block is added successfully to the blockchain network.

4 Conclusion and Future Scope

By understanding the characteristics, features, and working model of the blockchain technology, we can certainly come to a conclusion that it is one of the big revolutionary technologies in today's world since it makes the life of a human very

simpler and safer. Blockchain technology has certain features that make it more useful resource for commercial applications and a possible disruptive source for established companies. This includes ledger immutability, data decentralization, privacy preservation, trust-less transaction allowance, process efficiency and sustainability, and the capacity to optimize multi-step procedures utilizing smart contracts. Ultimately, blockchain can be said to be the rising technology that uses cryptography to change the world by enhancing the digital transactions that can happen anytime and anywhere in the world.

Researchers are also trying hard to overcome the challenges that are existing in this technology, and several literature reviews are done and its increasing day by day followed by the publications in this blockchain technology. While blockchain principles are used for different purposes, there are few numerical flaws. Blockchain uses lots of energy to create a new block, and a lot more energy to create the hash value for every blocks [12]. The scope and existence of this blockchain technology will be tremendously inevitable in tomorrow's digital world if these drawbacks are sorted by upcoming researchers.

References

1. Peck M (2018) Understanding blockchain technology: abstracting the blockchain. IEEE
2. Yaga D, Mell P, Roby N, Scarfone K (2019) Blockchain technology overview. arXiv preprint arXiv:1906.11078
3. Michael J, Cohn ALAN, Butcher JR (2018) Blockchain technology. Journal 1(7)
4. Viriyasitavat W, Hoonsopon D (2019) Blockchain characteristics and consensus in modern business processes. J Ind Inf Integr 13:32–39
5. Zheng Z, Xie S, Dai H, Chen X, Wang H (2017) An overview of blockchain technology: architecture, consensus, and future trends. In: *2017 IEEE international congress on big data (BigData congress)*, IEEE, pp 557–564
6. https://rubygarage.org/blog/how-blockchain-works
7. Monrat AA, Schelén O, Andersson K (2019) A survey of blockchain from the perspectives of applications, challenges, and opportunities. IEEE Access 7:117134–117151
8. Andreev RA, Andreeva PA, Krotov LN, Krotova EL (2018) Review of blockchain technology: types of blockchain and their application. Intellekt Sist Proizv 16(1):11–14
9. Madavi D (2019) A comprehensive study on blockchain technology. Int Res J Eng Technol 6(1):1765–1770
10. Soni S, Bhushan B (2019) A comprehensive survey on blockchain: working, security analysis, privacy threats and potential applications. In: 2019 2nd international conference on intelligent computing, instrumentation and control technologies (ICICICT) vol 1. IEEE, pp 922–926
11. https://intellipaat.com/blog/tutorial/blockchain-tutorial/how-does-blockchain-work/
12. Efanov D, Roschin P (2018) The all-pervasiveness of the blockchain technology. Proc Comput Sci 123:116–121

Topic Modelling-Based Approach for Clustering Legal Documents

Aayush Halgekar, Ashish Rao, Dhruvi Khankhoje, Ishaan Khetan, and Kiran Bhowmick

Abstract The justice system has been institutionalized around the world for a long time, increasing the number of resources available for and in this field. The colossal increase in dependency on the World Wide Web is commensurate to the increase of digitization of documents. Along with this growth has come the need for accelerated knowledge management—automated aid in organizing, analysing, retrieving and presenting content in a useful and distributed manner. For a fair, cogent and strong legal case to be built, the individual fighting the case must have access to case documents from not only several years ago but also a few months ago. Any particular part of any of these various cases that received a verdict in the previous years could be beneficial to the individual's case. Considering all these factors, it is evident to develop a search engine for legal documents which will provide the user with all the relevant documents it requires. Moreover, unlike widely accessible documents on the Internet, where search and categorization services are generally free, the legal profession is still largely a fee-for-service field that makes the quality (in terms of performance metrics used such as precision and recall) a key difference in services provided. This paper proposes a unique approach to cluster these documents using the mini batch k-means algorithm on dimensionally reduced sentence embeddings generated with the use of DistilBERT and UMAP. The proposed approach has been compared to state-of-the-art topic modelling and clustering approaches and has outperformed them.

A. Halgekar (✉) · A. Rao · D. Khankhoje · I. Khetan · K. Bhowmick
Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
e-mail: aayush.halgekar@svkmmumbai.onmicrosoft.com

A. Rao
e-mail: ashish.rao@svkmmumbai.onmicrosoft.com

D. Khankhoje
e-mail: dhruvi.khankhoje@svkmmumbai.onmicrosoft.com

I. Khetan
e-mail: ishaan.khetan@svkmmumbai.onmicrosoft.com

K. Bhowmick
e-mail: kiran.bhowmick@djsce.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_17

Keywords Document clustering · Sentence embeddings · Mini batch K-means · Topic modelling · UMAP

1 Introduction

The Internet contains a large amount of legal documents and whilst it facilitates the storage of documents and other data digitally, protecting them from physical damage and loss of information, it has its unique set of problems. One such problem being finding related articles or proceedings. In addition to the previous statement, the presence of myriad documents and resources has engendered a need for a search engine to enable discovery of the required documents in an optimized manner.

The legal domain is one such domain where the legal proceedings and articles have been digitized due to the increasing number of documents produced every day. The common law system is one of the most widely used legal systems in the world, and it is used in nations such as Australia, India, the United Kingdom, and the United States. The common law system's defining feature is that it places a high value on previous cases. It follows the idea of Stare decisis, which states that comparable facts or circumstances should result in similar outcomes. Simultaneously multiple cases are fought and could be setting a precedent for one another [1]. There are innumerable cases that have already been shut which still might play a small but extremely relevant role in an ongoing case. Additionally, there is also a possibility of a past case having an immense impact, but due to the disorganization and unshattered availability of these documents, they get missed. These problems indicate the urgent need for a way to group documents related to a topic and organize them better in order to be stored and retrieved efficiently. Considering all these characteristics and circumstances, a clustering of legal documents will enable organized, well-timed and quick search of these documents. Clustering is an unsupervised machine learning approach where one draws inferences from datasets consisting of input data without labelled responses. Generally, it is used as a process to find meaningful structure, explanatory underlying processes, generative features, and groupings inherent in a set of examples. By grouping related pages into one category, clustering is used to assist users in navigating, summarizing, organizing, and improving the search results of the huge amounts of textual data.

In the context of this paper, a clustering algorithm will serve to group topically similar legal documents together. This will not only help the legal community for knowledge management but also any individual who fails to understand the jargons of the legal community in obtaining the required information by browsing through similar articles and not having to explicitly search for them. To allow the legal community to retain both their broad coverage and productivity in their practice of law it is essential to adequately cluster legal documents and facilitate their organized access.

The proposed approach aims to create clusters that are internally cohesive but are extremely different from other clusters. The methodology of clustering text documents is to use their term frequency value to group the documents having akin terms in one cluster. The terms act as a measure of topic or the contents of a document and accordingly assign them to a cluster. Furthermore, with the recent advancements in the field of natural language processing, context-aware approaches which not only take into consideration the mere words present in a document but also their relationship with other words in a sentence have also been explored in this paper.

The key contribution of this work involves successfully developing a topic modelling-based clustering approach for the grouping of legal documents to create pure and cohesive clusters that outperform existing topic modelling and clustering models.

The paper discusses previous and related work in Sect. 2, the proposed methodology and steps employed during implementation of the proposed system in Sect. 3. Section 4 explains the dataset used, its analysis and comparative modelling and Sect. 5 discusses the results of various existing models and the proposed model.

2 Related Work

Document clustering algorithms have traditionally relied on algorithms used to perform unsupervised clustering on numerical or tabular data. The main premise behind such algorithms is the use of vectorization of the documents present and applying a traditional clustering algorithm. Authors of [2] work on improving the traditional k-means algorithm by dimensionality reduction which helps to make it faster and efficient. On the other hand, Chaudhary and Kshirsagar [3] focuses on the Agglomerative clustering algorithm which is another traditional clustering algorithm employed usually for numerical data.

In [4–6], the focus is on graph-based solutions to solve the problem of logical relevance during document clustering. In [4], the paper uses a novel graph approach for soft clustering. Authors of [5, 6] employ a graph-based solution based on background knowledge in the form of ontologies and along with it, k-means clustering.

Some other approaches which try to subsume a context-aware component in them are based on the usage of different types of similarity metrics. The idea behind these approaches is to check for the similarity of a document with the other documents present in the corpus. Traditionally, similarity metrics such as the Euclidean distance or the Manhattan distance, have been used. Though these similarity metrics may be helpful in providing the similarity between numerical values, they fail to capture the similarity between similar words. In order to tackle this problem, various methods have made the use of the cosine similarity of the vectors created on the basis of words in the document. However, in [7], the authors demonstrated that the use of cosine similarity alone is insufficient as it does not take into consideration the difference in magnitude of two vectors. In order to tackle this problem, they introduced a new similarity metric which provides a more accurate similarity prediction as compared

to traditional similarity metrics. In [8], another similarity metric has been introduced known as the ordered weighted average (OWA) in order to provide a similarity value.

In their paper, authors of [9] talk about the possible integration of the concept of topic modelling with that of document clustering. Their paper proposes a new model based on the LDA model for topic modelling. Based on the topics generated by the model, the clustering of documents takes place.

In order to provide further advanced solutions to the defined problem, Subramani et al. [10] proposes a novel approach for using artificial neural networks for the task of topic modelling for clustering documents. In [11], Glove word embeddings and DBSCAN algorithm for semantic document clustering are used. The authors of [12] compare different embeddings for the task of document clustering and topic modelling and compare results with the traditional LDA model as well. Sentence embeddings are paired with a dimensionality reduction technique followed by a density-based clustering approach in [13]. Authors of [14] propose a modification on [13] by adding c-TF-IDF which helps to extract the main words and understand the topics of the generated clusters. This paper proposes a new methodology to cluster legal documents, which performs better than existing architectures.

3 Proposed Methodology

This section describes the methodology implemented. Each section addresses a component in the overall procedure followed. The proposed methodology is illustrated in Fig. 1.

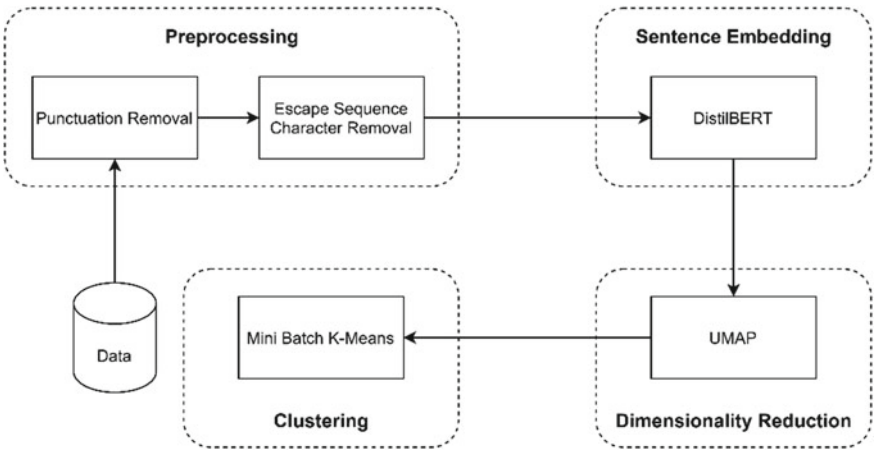


Fig. 1 Proposed methodology workflow

3.1 Preprocessing

The first step of the proposed methodology is the preprocessing of the documents. For preprocessing, a pipeline involving punctuation removal and escape sequence character removal is performed. Stopwords have not been removed in order to preserve the relationship of words. Additionally, lemmatization or stemming have been avoided in order to preserve the lexical structure of the data.

3.2 Sentence Embeddings

The pre-processed documents are then embedded into numeric values using a Sentence Transformer. The DistilBERT language model is used under the hood by the Sentence Transformer to generate the sentence embeddings. The distillation process in DistilBERT provides a quicker and lighter alternative to the traditional BERT model. DistilBERT is well suited for the corpus used as it is distilled on huge batches and takes advantage of gradient accumulation by utilizing dynamic masking and forgoing the next sentence prediction goal to provide better results as compared to BERT-base architecture [15].

3.3 Dimensionality Reduction

The embedded documents are spread across various dimensions which makes it more difficult to cluster. Hence, the next step in the proposed methodology is to perform dimensionality reduction. The Uniform Manifold Approximation and Projection or UMAP [16] is used to reduce the dimension of the embeddings to 5 dimensions. Since the document vectors in high dimensional spaces are very sparse, dimension reduction might help in finding dense areas. It is chosen for dimensionality reduction in this paper, as it preserves local and global structure, and is able to scale to very large datasets [13]. The most important hyper-parameter out of the several UMAP has to determine its way of performing dimension reduction is the number of nearest neighbours. This parameter is responsible for maintaining balance between preserving global structure versus local structure. More emphasis is put on global over local structure preservation by larger values.

3.4 Clustering

The dimensionally reduced embeddings are then clustered using the mini batch k-means algorithm [17], which is a variant of the traditional k-means algorithm. The

algorithm employs the use of a series of iterations to continuously update the centroids of the clusters to get a dynamic cluster formation. In each new iteration, the document is assigned a cluster based on the current position of the centroids. Once the document is assigned to a cluster the position of the centroid is recalculated based on gradient descent which is faster than the traditional K-means algorithm.

It uses mini-batches to reduce the computation time in large datasets. Additionally, it attempts to optimize the results of the clustering. To achieve this, the mini batch k-means takes mini-batches as an input, which are subsets of the whole dataset, randomly.

4 Experimentation

4.1 Dataset Description

The dataset used in the paper is an open-source labelled dataset, which was used in [18] and made available on Open-Source Framework [19]. This corpus contains of legal cases including 58.16% taken from the United States Supreme Court. These cases are labelled into various categories like, Construction, Health Law, Family Law, Military and several more.

This paper uses 3 categories, namely, Military, Banking and Environment from the corpus for clustering of the legal documents belonging to these topics. Each of these categories have 218, 380, 881 documents, respectively, totalling to 1479 documents.

4.2 Dataset Analysis

Figure 2 depicts the number of documents in the original 3 topic groups. As observed the original documents were non-uniformly distributed with a certain group having clear dominance.

Figure 3 depicts the percentage distribution of the documents in the original 3 groups.

As visible in the above diagrams, the dataset was skewed and had a higher proportion of documents belonging to the environment group. The created clusters also have a similar distribution. The competency of the model in such a skewed dataset implies that the model will have accurate responses to balanced datasets as well.

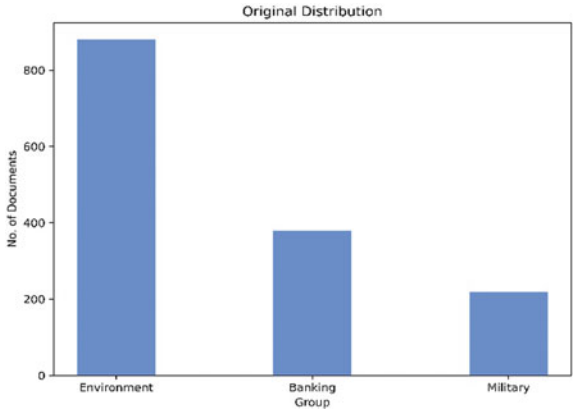


Fig. 2 Original distribution of documents

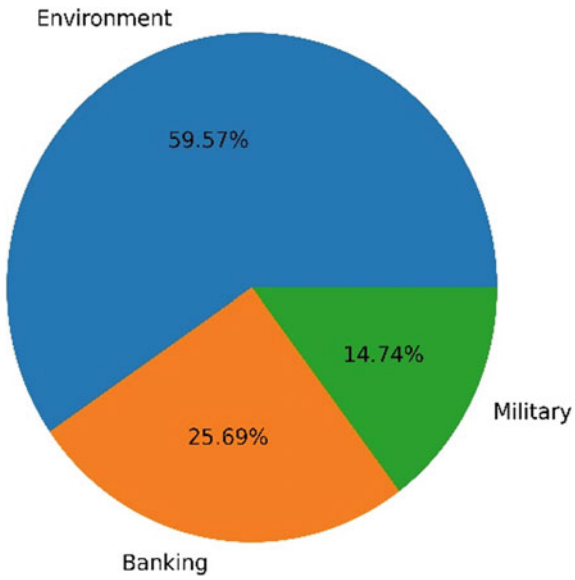


Fig. 3 Original percentage distribution of documents

4.3 Comparative Modelling

The proposed methodology was tested against state-of-the-art approaches for document clustering. The python libraries, genism [20] and Scikit-learn [21] were used for modelling. The proposed methodology involves sentence embeddings followed by UMAP and then mini batch k-means. The other models used are—Latent Dirichlet Allocation and BERTopic [14]. Another pipeline was tested which is, sentence embeddings followed by UMAP and then, clustering using k-means. This pipeline

was added in order to demonstrate the effectiveness of the mini batch k-means algorithm. Lastly, sentence embeddings without UMAP was also tested. This helped to show the significance of dimensionality reduction on the results.

5 Results and Discussion

5.1 Performance Metrics

In order to evaluate the performance of the models used for the task of document clustering, certain performance metrics have been selected. Since the dataset provides labels denoting the topic of the documents, it enables the usage of certain performance metrics such as purity, precision, recall and F1-score which are dependent on the presence of labels.

Since there are multiple topics which are being evaluated, there will be multiple values of precision, recall and F1-score as they depend on the classes or topics in the data. In order to analyses the results better, the macro average of these metrics is compared. Macro average is selected in preference to macro weighted average due to the imbalance in the topics present in the data.

Table 1 shows the results recorded during experimentation with each model for results for purity, homogeneity and completeness. Table 2 shows the results for precision, recall and F1-score for each model.

Table 1 Results for purity, homogeneity and completeness for each model

Model	Purity	Homogeneity	Completeness
DistilBert + MiniBatch K-Means	0.66	0.27	0.25
LDA	0.76	0.37	0.32
BERTopic	0.84	0.44	0.47
DistilBERT + UMAP + K-Means	0.77	0.53	0.48
DistilBERT + UMAP + Mini Batch K-Means	**0.85**	**0.53**	**0.49**

Table 2 Results for precision, recall and F1-score for each model

Model	Precision	Recall	F1-score
DistilBert + Mini Batch K-means	0.48	0.27	0.40
LDA	0.19	0.28	0.21
BERTopic	0.62	0.65	0.67
DistilBERT + UMAP + K-Means	0.75	0.54	0.62
DistilBERT + UMAP + Mini Batch K-Means	**0.78**	**0.66**	**0.71**

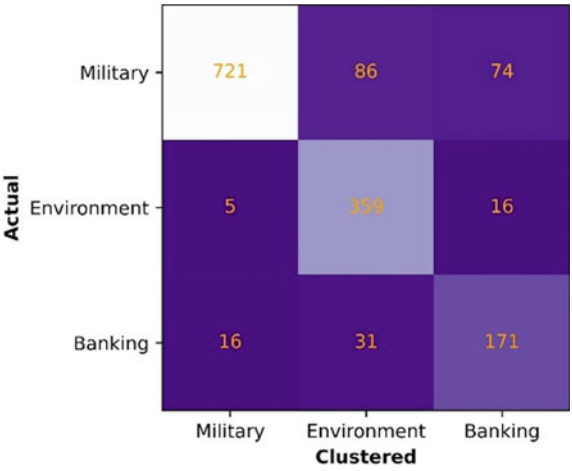


Fig. 4 Confusion matrix of proposed model

5.2 Result Analysis

On analysing the performance of the models using the performance metrics mentioned, it can be observed that the sentence embeddings + UMAP + Mini Batch K-Means architecture produces the best-case results and surpasses state-of-the-art models in the task of topic modelling and clustering documents according to the topic. The results indicate the importance of dimensionality reduction step on comparing the best-case results with the sentence embeddings + Mini Batch K-Means architecture. It can also be observed that modifying the traditional k-means clustering algorithm also provides better results. The best-case results can be further analysed with the help of a confusion matrix which helps to understand the clusters better.

Figure 4 shows the confusion matrix of the results produced by the sentence embeddings + UMAP + Mini Batch K-Means architecture.

As it can be seen, the confusion matrix shows the number of documents belonging to each topic and the actual and the clustered number of documents. It can be observed that the topic 'Military' has 721 documents which belong to a single cluster which is around 81.83% of the documents belonging to the topic. Similarly, the topic 'Environment' has 359 documents and the topic 'Banking' has 171 documents belonging to their corresponding single clusters which is around 94.47% and 78.44% of the documents, respectively. It can be concluded that the architecture performs well for all three topics. Along with the external measures for evaluation, internal metrics such as purity, homogeneity and completeness show that the architecture produces clusters which are not very overlapping and distinctive as compared to the clusters formed by other architectures.

The proposed methodology has a minor limitation in the nature of its results. It has been observed that multiple topics can be ascribed to certain legal documents. Due to the use of mini batch k-means in clustering stage, the proposed model does not take into factor the possibility of one document belonging to multiple clusters.

6 Conclusion and Future Work

The proposed model for legal document clustering successfully employs a context-aware approach with the use of the mini batch k-means model. With the help of this model, topic modelling is carried out and the words with high influence in each topic indicate a logical relevance to each other in the legal context. The documents are categorized into clusters based on the topics outputted by the model. The performance of the system is encouraging because it produces results better than state-of-the-art topic modelling and clustering algorithms. Through the analysis of the corpus, it is evident that there lies an imbalance in dataset, implying that the system can even perform satisfactorily in such conditions.

Clustering of legal documents still continues to be difficult problem because of the diverse nature of these documents. There are several legal cases whose documentation is multi-topical in nature, insinuating that they could belong to more than one clusters. Future work can be done in researching on soft clustering algorithms and expanding the scope of legal document clustering to allow clusters to accurately belong to multiple clusters.

References

1. Mandal A, Chaki R, Saha S et al (2017) Measuring similarity among legal court case documents. In: ACM international conference proceeding series, pp 1–9. https://doi.org/10.1145/3140107.3140119
2. Wu G, Fu E, Lin H, Wang L (2016) An improved K-means algorithm for document clustering. In: Proceedings—2015 international conference on computer science and mechanical automation, CSMA 2015, Institute of Electrical and Electronics Engineers Inc., pp 65–69
3. Chaudhary G, Kshirsagar M (2021) Enhanced text clustering approach using hierarchical agglomerative clustering with principal components analysis to design document recommendation system. Research transcripts in computer, electrical and electronics engineering, vol 2
4. Dörpinghaus J, Schaaf S, Jacobs M (2018) Soft document clustering using a novel graph covering approach. BioData Min 11(1):1–20. https://doi.org/10.1186/S13040-018-0172-X
5. Hotho A, Maedche A, Staab S (PDF) Ontology-based text document clustering (2002)
6. Svadas T, Jha J (2015) Document cluster mining on text documents. Int J Comput Sci Mobile Comput 778–782
7. Diallo B, Hu J, Li T et al (2021) Multi-view document clustering based on geometrical similarity measurement. Int J Mach Learn Cybern 1–13. https://doi.org/10.1007/s13042-021-01295-8
8. Wagh RS, Anand D (2020) Legal document similarity: a multicriteria decision-making perspective. PeerJ Comput Sci 2020:1–20. https://doi.org/10.7717/peerj-cs.26
9. Xie P, Xing EP (2013) Integrating document clustering and topic modeling
10. Subramani S, Sridhar V, Shetty K (2018) A novel approach of neural topic modelling for document clustering. In: Proceedings of the 2018 IEEE symposium series on computational intelligence, SSCI 2018, vol 28, pp 2169–2173. https://doi.org/10.1109/SSCI.2018.8628912
11. Mohammed SM, Jacksi K, Zeebaree SRM (2020) Glove word embedding and DBSCAN algorithms for semantic document clustering. In: 3rd international conference on advanced science and engineering, ICOASE 2020, pp 211–216. https://doi.org/10.1109/ICOASE51841.2020.9436540

12. Curiskis SA, Drake B, Osborn TR, Kennedy PJ (2020) An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit. Inf Process Manage 57(2):102034. https://doi.org/10.1016/J.IPM.2019.04.002
13. Angelov D (2020) TOP2VEC: distributed representations of topics
14. Grootendorst M, Reimers N (2021) MaartenGr/BERTopic: v0.9. https://doi.org/10.5281/ZENODO.5168575
15. Sanh V, Debut L, Chaumond J, Wolf T (2019) DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter
16. Mcinnes L, Healy J, Melville J (2020) UMAP: uniform manifold approximation and projection for dimension reduction
17. Sculley D (2010) Web-scale k-means clustering. In: Proceedings of the 19th international conference on World Wide Web, WWW' 10:1177–1178. https://doi.org/10.1145/1772690.1772862
18. Sugathadasa K, Ayesha B, de Silva N et al (2017) Synergistic union of Word2Vec and lexicon for domain specific semantic similarity
19. de Silva N, Ayesha B (2019) SigmaLaw—large legal text corpus and word embeddings. https://osf.io/qvg8s/. Accessed 11 August 2021
20. Rehurek R, Rehurek R, Sojka P (2010) Software framework for topic modelling with large corpora. In: Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks, pp 45–50
21. Pedregosa F, Varoquaux G, Gramfort A et al (2011) Scikit-learn: machine learning in python. J Mach Learn Res 12:2825–2830

Traffic Sign Detection and Recognition

Preeti S. Pillai, Bhagyashree Kinnal, Vishal Pattanashetty, and Nalini C. Iyer

Abstract Surge in the amount of automobiles on street imposes the consumption of automatic systems for driver aid. These structures form significant instruments of self-driving automobiles as well. Traffic Sign Recognition remains such an automatic structure which affords the relative responsiveness aimed at self-driving automobile. In this work we are able to perceive and identify traffic signs in video classifications detailed by an onboard automobile camera. Traffic Sign Recognition (TSR) stands used to control traffic signs, inform a driver and facility or proscribe definite actions. A debauched real-time and vigorous instinctive traffic sign finding and recognition can upkeep and disburden the driver and ominously upsurge heavy protection and ease. Instinctive recognition of traffic signs is also important for automated intellectual driving automobile or driver backing structures. This paper presents a study to identify traffic sign via OpenCV procedure and also convert the detected sign into text and audio signal. The pictures are mined, perceived and recognized by preprocessing through numerous image processing methods. At that time, the phases are accomplished to identify and identify the traffic sign arrangements. The structure is trained and endorsed to find the finest network architecture. Aimed at the network exercise and assessment we have generated a dataset containing of 1012 images of 8 diverse classes. The tentative results demonstrate the exceedingly accurate groupings of traffic sign patterns with composite contextual images and the computational price of the planned system. Though, numerous features make the road sign recognition tricky and problematic such as lighting state changes, occlusion of signs due to hitches, distortion of signs, gesture blur in video images.

Keywords Self-driving cars · Traffic signs · Machine learning · Embedded systems · Audio and text signals

P. S. Pillai (✉) · B. Kinnal · V. Pattanashetty · N. C. Iyer
K.L.E. Technological University, Hubli, India
e-mail: preeti_pillai@bvb.edu; preeti_pillai@kletech.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_18

1 Introduction

Driver assisting systems are used to evade accidents are apt added and more vital as the amount of vehicles on street is on an exponential increase. Advanced driver assistance systems (ADAS) are successfully used in automobiles for providing road keep assist; traffic sign aid systems, etc. [1]. Traffic sign detection is the central part factor of traffic sign aid system used for providing appropriate commands and warnings to the drivers concerning traffic limitations and proper information. At night time or in ghastly weather conditions, traffic signs are stiff to make out properly as well as the drivers are simply exaggerated by headlights of approaching vehicles. These conditions possibly will direct to traffic accidents plus grave injuries. A vision-based highway, street sign recognition structure is therefore enviable to grasp the awareness of a driver to keep away from traffic hazards [2]. A traffic sign detection algorithm frequently consists of two modules: the discovery module and the categorization module. The discovery module receive metaphors from the camera and finds out every the regions in the metaphors that may include traffic signs; then the categorization module determine the grouping of traffic sign in each section. The data given by the traffic signs is determined in their visual properties: color, form, and pictogram. As a result, the recognition and the detection modules are based on color and shape cues of traffic signs [3]. This provide the growth of more intelligent vehicles that possibly will sense, comprehend and pertain the significance of each traffic sign in the street by giving the opportunity to perceive any signs of each input image by means of Image processing and convolution neural networks (CNN).

1.1 Motivation

In order to solve the concerns over transportation and human safety, traffic sign detection and recognition has been introduced. Single state detectors are suffering from less accuracy and have difficulty in detecting small objects. Traffic signs which are far from the vehicle would be appear smaller in the image so the single stage detectors are not considered suitable for the detection and recognition. In the recent advancements made with respect to accuracy and the small object detection issues have been tackled. This makes the worth studying the performance of the VGG 16 CNN model. This paper presents a study on the usage of the VGG 16 CNN model as used in the traffic sign detection and recognition system.

2 Literature Survey

With the incorporation of electronics into the automobiles these days the convenience and user experience has drastically increased. Due to this scenario the traffic

environment on road is more complicated. Governments have put forth many rules and regulations to manage traffic on road and also to avoid unpleasant situations like accidents. Many a times the driver due to manual errors may miss the signals which lead to traffic problems and accidents. Traffic sign detection in modern cars is done through advanced concepts like machine learning [4]. A lot of work is being done to make the algorithms more accurate and efficient. A paper based on Gaussian kernel using support vector machine detects speed limits on the road. Here the images are taken from a database and it has 88.9% success rate. Feature used for detection and recognition is shapes of the signals and classification is done using CNN. The ROI in this paper is found by shape and color of the traffic signal. The procedure is immune to motion blur and also performs well in condition where there is low light [5].

Another paper, traffic sign detection is done by three main steps. The first one is being preprocessing where image enhancement is done by removing the noise. The second step being detection generation of signs is done. Color is the basis of image segmentation. The third and final step is recognition where the ROI is extracted. In the third stage classification of the images extracted is done and finally recognition with accuracy of 94% is done using CNN. In another paper, Nanal et al. speaks about zonal and speed controlling using a controller. The controller has two parts, of which one is a transmitter for to send out status of zones and second is the control unit with a display. The RF transmitters on the traffic signals in the particular zones send out traffic signal to the controlling unit in the vehicle and then are displayed on the display unit. The driver can adjust the speed of the vehicle according to the zone speed information received [6].

In the work presented by Sun et al., a classification model is discussed. It classifies traffic signals by first color segmentation and recognition then classification. Model works on the images taken from vehicles in motion. A shadow highlight invariant algorithm is chosen for color segmentation and its 97% accurate. Further, fuzzy shape recognition is used for recognition. Support vector machine is used as a classifier [7].

In the work proposed by Rajendran et al., Yolo V3 architecture is used for detection and classification is done through CNN. The detection is achieved at a speed of 10 frames per second with high MAP performance of 92%. The detection model is able to categorize and put the bounding box for most of the signals. Very few signs could not be localized properly by the model. The proposed detector is able to detect almost all categories of traffic sign and good progress accurate bonding box off most of the detected signs. Few very small traffic sign could not be accurately localized and also there were little false positive detection. Traffic sign classifier used in simple in architecture with very high accuracy. As future step simultaneous detection and classification of traffic sign using single stage detectors could be explored this method could help avoiding the use of an additional traffic sign classification network [8–10].

2.1 Organization of the Paper

In this section, we discuss about the state of art approaches used for traffic sign detections and recognitions. Section 3 discusses about the Implementation details includes the specifications of traffic sign recognition and conversion and algorithms. Section 4 discusses about the obtained results and its discussions. Section 5 concludes the work.

3 Framework for Traffic Sign Detection and Recognition

The proposed framework for traffic sign detection and recognition is shown in Fig. 1. The framework proposed is divided into two parts

- Traffic signal detection
- Traffic signal recognition.

Video is taken as input and they are converted into frames. Then images are created using image extraction blocks. Preprocessing of images, i.e., blurring of image, resizing of image is done. Color and shape are the major part of detection. As most of the Indian traffic signs are in red and white color, color space conversion is done on the preprocessed image which translates red color to white after which Corner detection of white part is done. Only corners which have rectangle shape, triangle shape and round shape are considered because traffic signals are mainly in three shapes. Then the signboard is detected. After detection of signboard traffic signal is recognized using CNN model and output is taken in the form of text and voice.

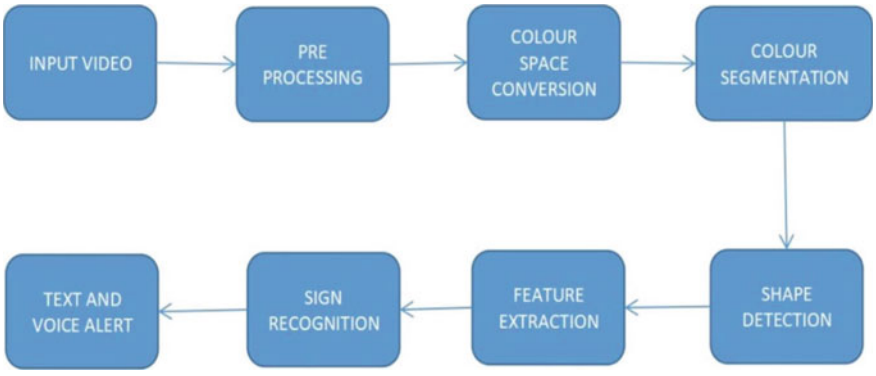


Fig. 1 Proposed system design for traffic sign detection and recognition

3.1 Traffic Sign Detection

For traffic signal detection we are using image processing as our main tool. The implementation is done in Python and the image processing tool used is OpenCV. The first step is to take input from the web camera or a video. This video is then converted to frames to detect traffic signals in a given video. Each and every image has been evaluated by using the OpenCV tool. First the input image is blurred for evenly distribution of color in the image.

The image is converted into binary format. In binary form images, white color resembles parts of the image with red color, and the rest part of the image which has non-red color resembles black color. All the corners of the white part in an image are computed. The corners are filtered to take the corners which have only rectangle shape, triangle shape or round shape, because we have traffic signal mainly in three shapes. This shaped corner is considered as one blob. These are the main steps which are carried out to find location of traffic signals in a given image which is called as traffic signal detection.

3.2 Traffic Sign Recognition

Traffic sign recognition is done using a trained CNN model. The data set containing 2000 + images is used for training and testing the CNN model. The dataset contains 2000 + images of 8 different traffic sign for training the CNN model. The classes used in the dataset are as U turn, Double bend sign, Road join, speed limit 40, speed limit 50, speed limit 60 and Side road left.

CNN Training: The first step in training is to group the images with similar size or dimensions for processing. The dimensions of the images in the constructed dataset varies from 16 * 16 * 3 to 128 * 128 * 3. Hence the images from the data sent cannot be directly used to train the CNN model. Also there is an imbalance in the dataset with respect to different classes. This might make our model biased toward higher images class, particularly in situations where the model is not sure of the correct prediction. The dataset considered have samples images from real world environment so that the model is trained to handle the real-time conditions.

Preprocessing of Data: In order to process the images all the images are compressed to single dimension. This compression process has to compress the image to such a size that it does not lose the information nor stretches it too much and loses the accuracy. Taking into consideration these factors the images are compressed to 32 $\times$ 32 $\times$ 3 dimension.

Model Architecture: The CNN model used is VGG 16 model shown in Fig. 3 and its layers in Fig. 2. It is one of the commonly used models for image processing now a day. This pre-trained model is firstly used for our image-based traffic sign dataset

Layers	Description
Input Layer	32x32x1 images
Convolution-1	Convolution and rectified linear activation (ReLU).
Pool-1	Max pooling.
Convoution-2	Convolution and rectified linear activation.
Pool-2	Max pooling.
Local-3	Fully connected layer with ReLU
Local-4	fully connected layer with ReLU
softmax	Classification result

Fig. 2 CNN layers

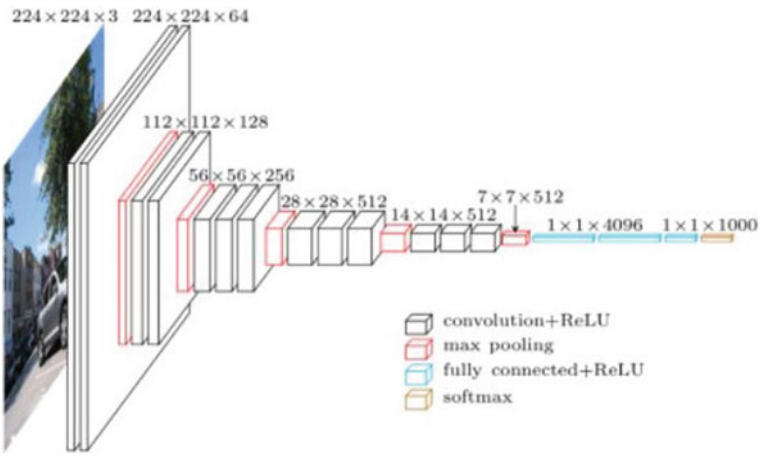


Fig. 3 VGG 16 model

which is then saved and used in recognition part of the detected images which are extracted from the videos (Fig. 3).

- The input layer has the input images which are represented as an array of pixels values.
- Next is the convolution layer which performs dot product between sub array of input and the kernel. The result of the dot product is a single value which would act as a pixel of the output image. This process is continued for the entire size of the input image and an output image of lesser dimension is generated.
- Next layer in the architecture is the RELU layer which applies an activation function max $(0, x)$ on the output image pixel values.
- After applying the activation function the output image is passed to the POOL layer which will down sample the width and height of an image resulting into reduced dimension image.
- Finally the fully-connected layer will compute the class scores for every image of the classification category.

Key Features of VGG 16

- Another name of VGG 16 is OxfordNet model, named after the Visual Geometry Group from Oxford.
- The model has 16 layers with associated weights.
- The model has only convolution layer and pooling layers in it.
- The size of the kernel used for convolution is 3 × 3 Kernel.
- The size of max pooling is 2 × 2.
- The model uses around 138 million parameters.
- Dataset used for training the model is ImageNet.
- The accuracy of the model is around 92.7%

The CNN model finally converts the input image pixel values to detected class. The RELU and POOL layers are implemented as constant functions hence no training takes place at these layers. The optimization function used to reduce the loss is gradient descent optimizer.

4 Results and Discussion

Recognizing traffic signs we will be using the saved CNN model which is trained using 8 different traffic signs which has the ability to recognize a given traffic sign. This recognizing part will take a blob, which has been identified in the image processing part. And we resize this blob into 32 * 32 sizes. This image will be fed into the CNN model for the recognition of the traffic signs. The recognized traffic sign is converted to text and voice alert and showed as an output to user.

4.1 Training Result

Accuracy of Training: 99.84%

Accuracy of Testing: 94.96%

Figure 4 shows the accuracy and loss graphs. The left graph shows the accuracy of training and testing. The model was trained for around 35 epochs. The right side of the image shows the training loss.

4.2 Traffic Sign Detection

Figure 5 shows the sample results of traffic sign detection. The image on the left shows result of “NO U TURN” traffic sign detected and a bounding box is drawn for visual representation. The image on the right shows result of “ROAD JOIN” traffic sign detected and a bounding box is drawn for visual representation.

```
------------------------------
Optimizer:  Adam
Learning rate:  1e-05
Epochs:  50
==============================
Trained in 7.82 minutes
Results at the end of training: acc=99.84%, test_acc=94.96%
```

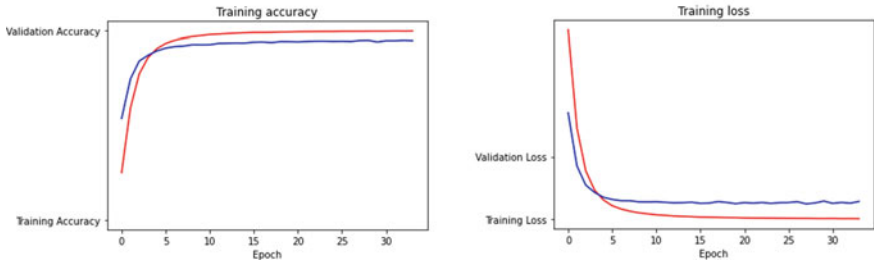


Fig. 4 Left graph shows the accuracy of training (red line) and testing (blue line) and the right graph shows the training loss

Fig. 5 Detection of "no u turn" sign board (left figure) and detection of "road join" sign board (right figure)

4.3 Traffic Sign Recognition

See Fig. 6.

5 Conclusion

In this paper, a work of detecting and recognizing traffic signs using the VGG 16 is discussed. The method uses both color and shape features for detection and further for recognition we used the VGG 16. We have achieved the results at the end of

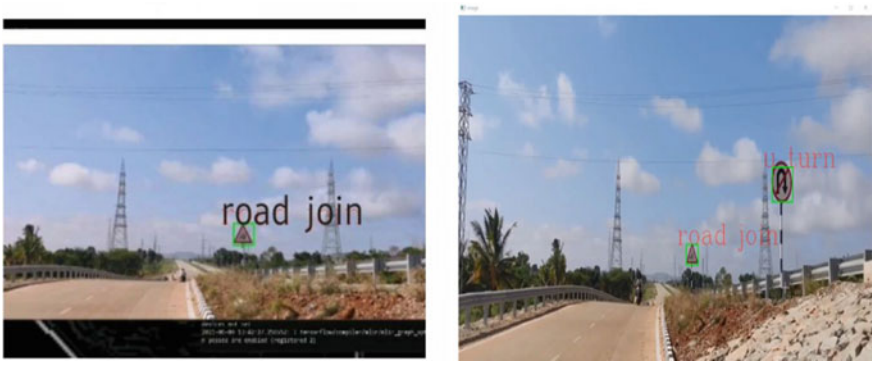


Fig. 6 Detection and recognition of "road join" sign board (left figure) and detection of "no u turn" sign board (right figure)

training, accuracy of training is 99.84 and accuracy of testing is 94.96%. The proposed detector is able to detect most of the traffic signs and could accurately put bonding boxes to most of the detected signs. The color and shape features-based detector and CNN-based classifier completes the pipeline of traffic sign detection and recognition. Finally a text and voice alert is given to alert the driver about the traffic signs.

References

1. Dhanakshirur RR, Pillai P, Tabib RA, Patil U, Mudenagudi U (2019) A framework for lane prediction on unstructured roads. In: Thampi S, Marques O, Krishnan S, Li KC, Ciuonzo D, Kolekar M (eds) Advances in signal processing and intelligent recognition systems. SIRS 2018. Communications in computer and information science, vol 968. Springer, Singapore. https://doi.org/10.1007/978-981-13-5758-9_16LNCS Homepage. http://www.springer.com/lncs. Last accessed 21 Nov 2016
2. Dhar P, Abedin MdZ, Biswas T, Datta A (2017) Traffic sign detection—a new approach and recognition using convolution neural network. In: IEEE region 10 humanitarian technology conference (R10-HTC) 2017 IEEE
3. Nanal A, Motwani P, Pawar P, Nirhale R (2019) Automatic detection of road signs to control vehicle speed. Int J Comput Appl 178:36–39. https://doi.org/10.5120/ijca2019918966
4. Iyer NC et al (2020) Millimeter-wave AWR1642 RADAR for obstacle detection: au-tonomous vehicles. In: Saini H, Singh R, Tariq Beg M, Sahambi J (eds) Innovations in electronics and communication engineering. Lecture notes in networks and systems, vol 107. Springer, Singapore. https://doi.org/10.1007/978-981-15-3172-9_10
5. Pattanashetty VB, Iyer NC, Viswanath HL (2018) Intellectual conveyance structure for travellers. In: Saeed K, Chaki N, Pati B, Bakshi S, Mohapatra D (eds) Progress in advanced computing and intelligent engineering. Advances in intelligent systems and computing, vol 564. Springer, Singapore. https://doi.org/10.1007/978-981-10-6875-1_37
6. Rajendran SP, Shine L, Pradeep R, Vijayaraghavan S (2019) Real-time traffic sign recognition using YOLOv3 based detector. In: 2019 10th international conference on computing, communication and networking technologies (ICCCNT), pp 1–7. https://doi.org/10.1109/ICCCNT45670.2019.8944890

7. Shraddha B, Subhas M, Nikita P, Preeti P, Bhagyashree K, Shet R, Jyoti P, Iyer N (2016) Model based learning of linear integrated circuit course: a practical approach. J Eng Educ Transformations. https://doi.org/10.16920/jeet/2016/v0i0/85549
8. Sun Y, Ge P, Liu D (2019) Traffic sign detection and recognition based on convolutional neural network. Chin Autom Congr (CAC) 2019:2851–2854. https://doi.org/10.1109/CAC48633.2019.8997240
9. Pattanashetty VB, Iyer NC, Dinkar A, Gudi S (2017) Self-coordinating bus route system to avoid bus bunching. In: Satapathy S, Bhateja V, Joshi A (eds) Proceedings of the international conference on data engineering and communication technology. Advances in intelligent systems and computing, vol 469. Springer, Singapore. https://doi.org/10.1007/978-981-10-1678-3_50
10. Pattanashetty VB, Iyer NC, Ibal P, Bagewadi A (2015) Smart driving assistance. IEEE Int Transp Electrification Conf (ITEC) 2015:1–4. https://doi.org/10.1109/ITEC-India.2015.7386870

Lung CT Scans Image Compression Using Innovative Transformation Approach

J. Vasanth Wason and A. Nagarajan

Abstract Doctors face difficulty in diagnosis of lung cancer due to the complex nature and clinical interrelations of computer diagnosed scan images. An image is an artifact which depicts visual perception. It contains large amount of data, hence requires more memory storage and causes incontinence for transmission over the limited bandwidth channel. By removing excessive redundant data, image compression technique reduces the amount of data required to represent an image and reduces cost of storage and transmission. Before compressing an image we are performing image transformation which consists of more number of zeros when compared to number of one's. Image compression minimizes the size in bytes of a graphic file without degrading the quality of the image. The main aim of compression is to reduce redundancy in stored or communicated data in turn increasing effective data density. In this research work normal c code is performed the transformation and compression is being performed to get the compressed image.

Keywords Coding redundancy · Inter pixel redundancy · Psycho visual redundancy · DCT · DPCM

1 Introduction

Dealing with medical images is not new in the field of medical research, the practitioners often use the MRI scans, and CT scans into consideration in order to know the patient's disease status. To understand these images requires expertise and time to extract the information from them [1]. When image are to be transmitted over the Internet it requires large storage space, requires high bandwidth, and it takes longer time for transmission. When such large images are to be transmitted they must be

J. Vasanth Wason (✉) · A. Nagarajan
Department of Computer Applications, Alagappa University, Karaikudi, India
e-mail: vasanthwasonj@gmail.com

A. Nagarajan
e-mail: nagarajana@alagappauniversity.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_19

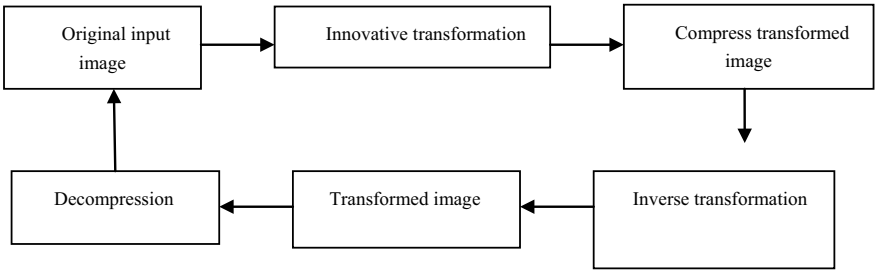


Fig. 1 Innovative transformation block diagram

represented with small number of bits hence image compression plays a very important role in transmission of an image which minimizes the storage requirement and speeding up across the low bandwidth channels [2].

The main aim of image compression is to reduce the number of redundant information without any loss of image quality. The images are stored in the form of pixels where pixels in turn are represented in the form of bits and these pixels are sampled at regular intervals which is called as sample point to obtain a digital image [3]. Each bit in the pixel consists of information which is stored numerically in the binary base. When the precision value is high then better will be the quality of an image. Initially the image is transformed using transformation technique then the transformation is applied prior to storage. The compression techniques are applied on the transformed image then the image obtained is a compressed transformed image [4]. In the compression technique the reduction operation is performed where it reduces the number of redundant information. The inverse transformation is applied on the compressed transformed image. In the decoding phase the decompression technique [5] is applied to the compressed transformed image to reconstruct the original image or an approximation of it (Fig. 1).

When compression technique are applied then only fewer bits are required to represent an image when compared to that of the original image hence compression techniques are used to facilitate storage capacity, the transmission bandwidth, the transmission time, which are required to transmit an image and the distribution of an image over the Internet, digital photography, remote sensing and medical imagery. The main characteristics of an image are that the neighboring pixels are co-related hence it contains the redundant information; hence it should be less core lance between the neighboring pixels.

Fundamental Components of Image Compression:

(a) Redundancy: Redundancy [5] is nothing but which consists of duplicate data with that of the original data.
(b) Irrelevancy: it is the omission of part of the signal data which cannot be noticed by a human eye when the signals are received by the signal receiver.

2 Image Transformation Process

When original image is given as an input for transformation function, then by applying transform we get another transformation mechanism image which is of same size [6]. Then when inverse transform is applied to the transformed image then the original image is obtained back [7]. When image transformation is performed it provides better image processing: The main characteristics of image is that it consists of high correlation between the neighboring pixels, hence it contains high redundant information hence transformation takes into account the long-range correlation in the space [7]. Conceptual insights in spatial frequency information. The images are represented in the form of pixels which in turn are represented in the form of bits, which are of different frequencies. Based on the frequency of information in the space smoothing moderate changes and fast changes are performed. Transformation is applicable in the area like diagnosing, enhancement and restoration (Fig. 2).

In image transformation it provide fast computation and alternative representation and sensing: The radiology images are given as an input and by using transformation technique the output image is a transformed data as a measurement for radiology images and then the inverse transformation is applied to recover the image [8]. Image transformation also provides efficient storage and transmission: Transformation is applied prior to storage. When images are to be transmitted over the Internet it requires large bandwidth, large storage capacity and requires more time to transmit hence transformation provide efficient storage and transmission. After the images are transformed the space consists of high energy compaction. Store/send the representation or contribution forms each basis.

Image Transformation Types

(1) Point to point transformation: The output value of a particular coordinate in an output image is specifically dependent only on one input value from the input image where it is not necessary to be of the same coordinate. It includes operations like rotations, scaling, Look-up table (LUT). It does not require an extra memory. They are called as neighborhood operators.

(2) Local to point transformation: the output value at a specific coordinate of an output image is mainly dependent on the input value in the neighborhood of that same coordinate.

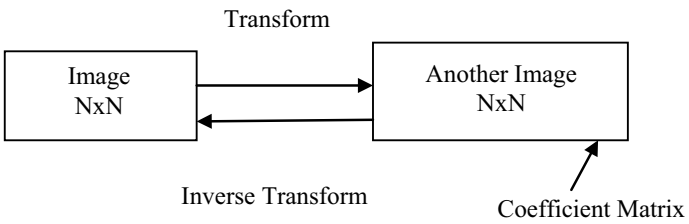


Fig. 2 Image transformation process

(3) Global to point transformation: the output value in an output image at a specific coordinate is completely dependent on all the values of an input image. When neighborhoods are considered then the complexity of an output image at a specific coordinate increases.

3 Background Study

Different examples which uses transformation technique [9]: Discrete cosine transform (DCT) which is based on linear transformation using blocks—There are various wavelet decomposition—Differential pulse code modulation (DPCM) which is based on prediction or residual formation—RGB to YcrCb which is based on color space transformation. Fractals which is based on model prediction or residual formation. DPCM is a prediction [10] mechanism, where it predicts the current pixel value (X) based on the combination of previously encoded pixels (A, B, C). It uses Huffman coding to encode the difference between the actual value and the predicted value ($-X$) [11]. The quantized difference is encoded in lossy DPCM. Lossless JPEG and 4.3 DPCM are based on DPCM. Properties which are followed by DPCM: (1) Low complexity (2) High quality (3) Low memory requirements DCT in DCT 8 * 8 block of Lung image is being transformed where each pixel value has been level-shifted by subtracting a value of 128 (Fig. 3).

Image representation with DCT: DCT uses an 8 * 8 block, where the coefficient can viewed as a weighting functions, when the coefficient is applied to 64 cosine basic function of different spatial frequencies, it will reconstruct the original block (Fig. 4).

DCT 8 * 8 image block: In 8 * 8 DCT, a single DCT coefficient will preserve the blocks energy (sum of squared amplitudes), by removing the pixel redundancy it packs the complete block energy into a single DCT coefficient (Fig. 5).

X (K, I) =

8	14	23	37	52	68	73	82
6	14	24	37	46	67	74	81
3	11	28	35	48	62	72	82
4	13	22	28	44	61	69	86
5	11	18	30	40	59	72	86
5	9	16	29	39	58	74	83
1	8	16	31	38	59	75	80
2	11	18	30	37	57	69	82

Fig. 3 8 * 8 DCT representations

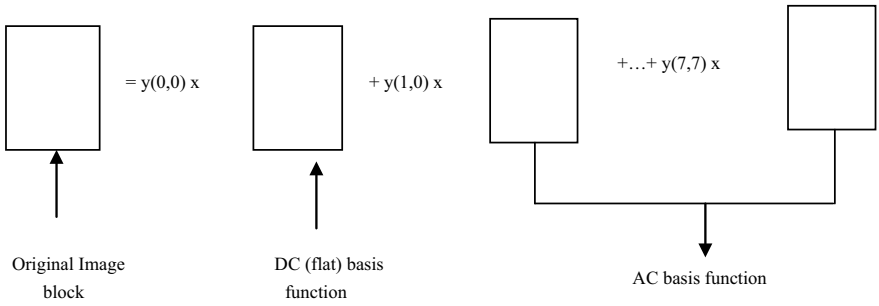


Fig. 4 Image representation with DCT

DCT coefficient

Y (U, V) =

327.5	-215.8	16.1	-10.7	-3.7	-1.5	4.2	-6.7
18.1	3.4	-9.9	3.7	0.5	-3.2	3.5	2.2
2.5	1.3	-5.4	2.8	-1.0	2.3	-1.6	-2.6
0.6	-2.5	3.0	5.0	1.8	2.2	-2.6	-1.4
0.3	1.6	3.4	0.0	2.5	-5.1	1.6	-0.7
-0.6	-1.8	-2.4	0.5	-0.4	-1.6	-0.1	2.1
0.9	1.6	-0.6	-0.7	2.1	-0.5	-0.9	2.8
0.6	-1.0	-2.9	-1.4	0.2	1.9	-0.6	0.7

Fig. 5 Representation of DCT coefficient

4 Proposed Algorithm

Algorithm for innovative transformation approach:

Step 1: start
Step 2: read input image to a file [original image]
Step 3: write a transformed image to an output file
Step 4: Initialize input array to zero
Step 5: Initialize output array to zero
Step 6: Adding 60 bytes of header to output file
Step 7: Character from the image file is stored in an array [characters from input image are read to file and they are stored in an array].
Step 8: Two pixel values are read from an input array

Step 9: Transformation process is performed [the first pixel value is considered as row value and second pixel value is divided by 8, the pixel value bit position is made as one in output array.
Step 10: Print output array.
Step 11: Close input file.
Step 12: Close output file.

8 bit Transformation: The 8 bit transformation we are considering original lung raw image where the image is read into a file. Different type of image requires different number of bytes for header, for lung raw image it requires 60 bytes of header. The content of the file are read to a dimensional array. The 8-bit format which of $28 = 256$ on an average we get $256/8 = 32$ best options. In an image it consist of 262,144 pixel and each pixel is represented in the form of bits. By considering the absolute function. $(p(x, y)-p(x, y-1))$ we obtain 0 to 31 best options which are of 5-bits and the magnitude of 0 to 31. Here an image of 512 * 512 original raw is considered and each pixel is of 8-bit representation and after transformation they are represented in 5-bit magnitude which consists of maximum of 1-ones or 2-ones or 3-ones.

After reading an image file into a two dimensional array, each pixel value is read into a variable value. The value is divided by 32 we obtain quotient and remainder. The remainder again it is divided by 4 we get a new quotient and remainder. When the power of first two quotients and the power of last remainder are being added we obtain a new value which consists of maximum of 1-ones or 2-ones or 3-ones and 5 zeros. This is stored in a new array in the form of binary representation, and this array is stored in an output file. This procedure is repeated for all the values of the pixel. The output file consists of information in the unreadable form hence by converting P5–P2 we can now read the output files. To view a transformed image we are using the GIMP software. The obtained image is a transformed image which consists of less number of 1 s when compared to that of the original image (Figs. 6, 7, 8 and 9).

Comparison of original lung raw image and 8-bit transformation image. The following tables are the comparison of number of 1 s and 0 s in each layer of original and transformed image (Figs. 10 and 11).

5 Conclusion

Compression is the technique which is applied to an image where the redundant and irrelevant information can be reduced. The neighboring pixels will have a correlation hence the redundancy. In the innovative transformation approach where the transformation such as 8-bit transformation. During each transformation the number of one's gradually starts decreasing compared to the previous transformation. Then the compression technique is applied where the compressed image consist of very few characters when compared to that of the original image. Compression process is applied to the inverse transformed image, where the size of the file is smaller when compared to that of the original image. Inverse transform mechanism is applied to the

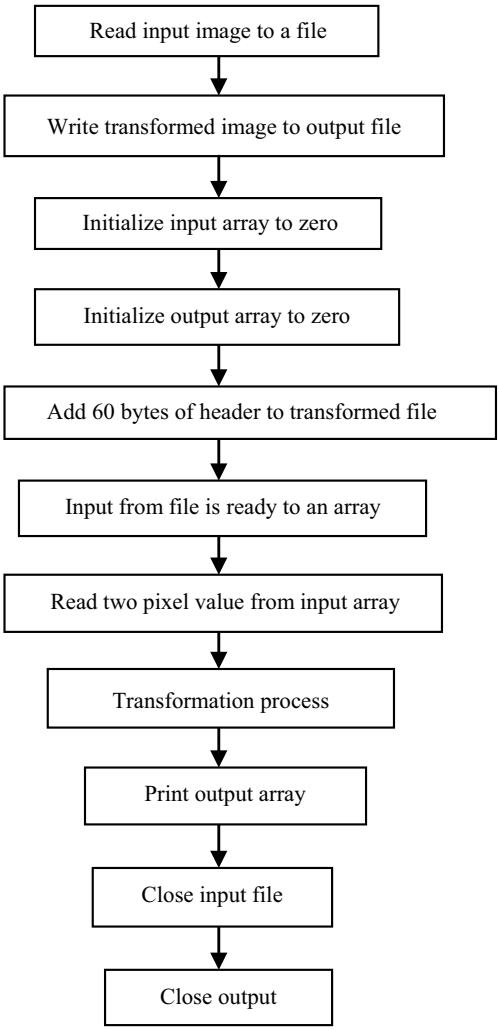
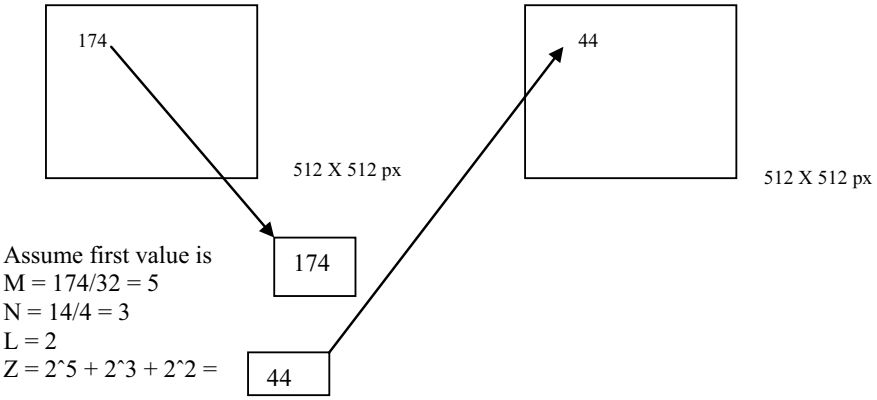


Fig. 6 Flow diagram of innovative transformation

compressed transformed image. Decoding is performed to the inverse compressed transformed image so that the resultant image is similar to that of the original image. Different type of transformation may be applied to the image by the quality of the image should not be reduced.

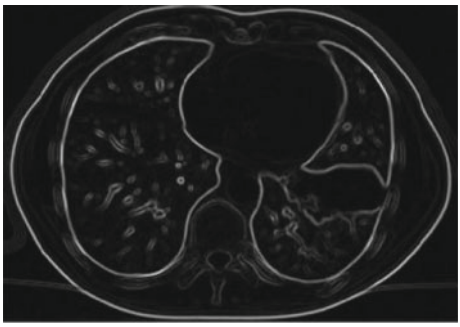


First pixel value is replaced with this new value which contains maximum 1, 2 or 3 ones.

Fig.7 8 bit transformation mechanism

Fig. 8 Lung raw image

Fig. 9 8 bit transformed lung image



Layer Number	Number of once	Number of zero	Total
0	1,31,304	1,30,840	2,62,144
1	1,31,244	1,30,930	2,62,144
2	1,30,658	1,31,486	2,62,144
3	1,30,774	1,31,370	2,62,144
4	1,31,797	1,30,347	2,62,144
5	1,32,745	1,29,399	2,62,144
6	1,10,124	1,52,020	2,62,144
7	1,34,000	1,28,144	2,62,144

Fig. 10 0 s and 1 s in original lung image

Layer Number	Number of once	Number of zero	Total
0	90,352	1,71,792	2,62,144
1	1,18,254	1,43,890	2,62,144
2	1,09,952	1,52,192	2,62,144
3	1,28,251	1,33,893	2,62,144
4	99,766	1,62,378	2,62,144
5	64,877	1,97,267	2,62,144
6	54,237	2,07,907	2,62,144
7	32,709	2,29,435	2,62,144

Fig. 11 0 s and 1 s in 8 bit transformed lung image

References

1. Sakthivel A, Nagarajan A (2018) Cancer detection methodology using fuzzy based classification techniques. Int J Eng Sci Res Technol 7(3):727–733. ISSN: 2277-9655
2. Basak P, Nath A (2017) Detection of different stages of lungs cancer in CT-scan images using image processing techniques. Int J Innov Res Comput Commun Eng 5(5):9708–9719
3. Karimi (2017) An intelligent system for quality measurement of golden bleached raisins using two comparative machine learning algorithms. Meas J Int Meas Confed 107:68–76
4. Abdillah B (2016) Image processing based detection of lung cancer on CT scan images. Asian J. 1–8. https://doi.org/10.1088/1742-6596/893/1/012063
5. Banu S (2015) 3D medical image compression: a review. Indian J Sci Technol 8(1)
6. Teymurazyan (2013) Properties of noise in positron emission tomography images reconstructed with filtered-back projection and row-action maximum likelihood algorithm. J Digit Imaging 26:447–456
7. Dodoo (2013) Trends in the utilization of medical imaging from 2003 to 2011: clinical encounters offer a complementary patient-centered focus. J Am Coll Radiol 10:507–512
8. Kumar V (2012) Design and implementation of 2D DCT/IDCT using Xilinx FPGA. Int J Eng Sci 4
9. Kusuma (2010) FPGA implementation of pipelined 2D-DCT and quantization architecture for JPEG image compression. In: International symposium on information technology (ITSim), Vol 1. IEEE
10. Talukder (2010) Haar wavelet based approach for image compression and quality assessment of compressed image. Available online: http://arxiv.org/abs/1010.4084

11. Kim KJ (2010) A comparison of three image fidelity metrics of different computational principles for JPEG2000 compressed abdomen CT images. IEEE Trans Med Imaging 1496–503

Effect of Data Compression on Cipher Text Aiming Secure and Improved Data Storage

M. K. Padmapriya and Pamela Vinitha Eric

Abstract Cryptography is a technique of protecting the data from an unauthorized access using encryption process. Encryption converts the plain text into cipher text which is in non-readable form. Past studies suggest that the size of cipher text is the major concern which hinders users from adopting encryption methods to secure the data. This work involves an experimental study on exploring the effect of data compression on an amino acid form of cipher text using dictionary-based methods as well as entropy coding methods without compromising the security. Compression ratio is measured for different file sizes. The results show that 47% storage savings is achieved using entropy coding method and 60% using dictionary-based coding method. Owing to this, storage efficiency is also doubled. The advantage of this method is thus convinced and provides an improved data storage.

Keywords Data compression · Encryption · Huffman coding · LZMA · Security

1 Introduction

In today's digital world, 2.5 quintillion bytes of data are getting generated every day. The rate at which the data are getting generated is going to increase due to the encroachment of technology. Data are the most valuable asset for any business organization in the era of big data. However, it is most important to protect the business data as data breaches are inevitable in today's context. This can be done using strong cryptography techniques, enhanced key management, and authentication policies. Cryptography methods can be symmetric, asymmetric encryption, or steganography methods. Encryption process converts the tangible form of a plaintext to a scrambled cipher text. Cipher text can be read only by the user having the authorized key.

M. K. Padmapriya (✉)
Research Scholar, New Horizon College of Engineering, Bengaluru, India
e-mail: padmapriyabhat@gmail.com

P. V. Eric
Professor, Presidency University, Bengaluru, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_20

Though cryptography protects data, it also increases the size of the encrypted data, thereby increasing the cost of storage and transmission. This is the reason most of the companies either will go for partial encryption or keep the data open which provides a room for cyber criminals [1]. Combining encryption and compression is a grim chore as there is less room for achieving high-compression ratio. On the surface, data compression and data encryption are discordant for the reason that encrypted data cannot be compressed by any known compression algorithms. Autosophy tree network data compression can combine encryption by using the library as the code key [2]. Data compression aims in reducing the size of a file which is very much essential in the era of digital world. The encryption is designed to realize the properties of compression algorithms [3]. In [4], the focus of the work was to incorporate chaos theory into arithmetic coding to achieve compression by maintaining confidentiality. It is ideal to choose encryption along with data compression to provide security and also storage optimization. Compression techniques like Huffman coding, arithmetic coding are combined with cryptographic security considerations in [5, 6]. Pipelining of data compression and encryption algorithms realized [7] which showed encryption mapping cannot improve the original data file in terms of compression. Small improvement is obtained for the pipeline of statistical algorithms and RSA encryption algorithm. Motivation behind this work is to determine the effect of data compression on cipher text toward storage optimization.

This work focusses on conducting an experimental study to determine the effect of compression on the cipher text which is in the form of amino acid codes. The objective of this work is to identify the compression ratio obtained using Huffman coding and LZMA compression methods on the amino acid cipher text which is a sequence of amino acid codes without an empty space. Execution time and compression ratio obtained using Huffman coding and LZMA compression methods are calculated and are compared with each other. This paper is organized as follows. Section 2 gives an overview of algorithms used, Sect. 3 presents results and discussion, and Sect. 4 is the conclusion.

2 Experimental Setup/Implementation/Methodology

This study uses a cipher text which is a result of DNA-based encryption. In DNA-based encryption, XOR operation is performed between the plain text and a DNA string obtained from a public database. The resulting string undergoes steganography operation where in random sequences of A, T, G, C codes are inserted into this intermediate cipher text. Further, this sequence is converted into amino acid codes by replacing DNA codons with their corresponding amino acid codes. Huffman coding and LZMA compression methods are used to determine the effect of data compression on the above cipher texts of different sizes. The test computer used was an Intel core i3@2 GHz machine with 4 GB RAM.

2.1 Huffman Coding

Huffman coding is an entropy coding method. It assigns flexible length code to the characters based on their number of occurrence in the input file. The most frequent character is assigned with smallest length code. Huffman coding prevents ambiguity by means of prefix coding, i.e., no character will have a code word which is a prefix of some other character code.

For a given set of alphabet $X = \{x_1, x_2, \ldots, x_n\}$ with a frequency distribution $f(x_i)$, the number of bits needed to encode a message of $\sum_{x=1}^{n} f(x_i)$ characters by using binary prefix code C is given by $B(C) = \sum_{x=1}^{n} f(x_i) \cdot L(c(x_i))$ where $c(x_i)$ is the code word for encoding x_i and $L(c(x_i))$ is the length of the code word $c(x_i)$.

Algorithm: Huffman Coding

Step 1: Select two characters p, q from alphabet X with the minimum frequencies and develop a sub-tree with these two letters as leaves. Let us label the root as r.
Step 2: Calculate frequency of r as $f(r) = f(p) + f(q)$.
Remove the characters p and q from the alphabet and insert r.
The new alphabet is now $X' = X \cup \{r\} - \{p, q\}$ where $|X'| = |X| - 1$.
Step 3: Repeat step 1 and 2 with the new alphabet X' until an alphabet with only one symbol is left. The resulting tree is a Huffman tree.
Step 4: Assign code 0 to the left edge and 1 to the right edge by traversing the tree from root
Step 5: Encode the input by replacing character with its code.

2.2 LZMA Compression

The Lempel–Ziv–Markov chain algorithm (LZMA) uses dictionary-based compression system. It finds matches using classy dictionary data structures and yields a creek of literal symbols and phrase references, which is encoded one bit at a time by the range encoder. Range encoding is implemented in binary and uses shift operations to divide integers rather than the slow division methods.

LZMA's model uses contexts specific to the bit fields in each representation of a literal or phrase. It is similar to generic byte-based model, but yields better compression as it avoids mixing of unrelated bits together within the same context. It uses larger dictionary sizes compared to other classic dictionary compressions.

LZMA uses a sliding window method with a search buffer on the left and a look-ahead buffer on the right. During compression, the string in the window moves from the right to the left. The search buffer is like a dictionary and contains the recently encoded string. The string to be encoded is present in look-ahead buffer. During compression, LZMA encoder searches the longest pattern in search buffer that matches the pattern in look-ahead buffer.

Basically, LZMA outputs three types. If nothing matches the look-ahead buffer, then a literal value in the range 0–255 is output. If a match is found, then a length

Table 1 Time taken for compression

S. No.	Cipher text size (KBs)	Huffman compression (s)	LZMA compression (s)
1	4527	28.16	7.99
2	13,534	223.49	26.85
3	17,939	404.12	38.84
4	26,821	898.18	57.27

and distance from the string in the look-ahead buffer to its match in the search buffer are output after being encoded by the range coder. Since the search buffer is large, a short 4-entry distance history array is used which contains the four most recent distances. In case, if the distance of the current match is found in one of the four distances in the distance array, then a length and a 2 bit index to the distance array are a output and are also encoded by the range encoder. Since LZMA uses very large dictionary for compression, uses hash chain or binary tree search methods to locate the matches [8].

3 Results and Discussion

Space and time efficiency are two most important factors for any algorithm. Compression ratio is the parameter used to measure the space efficiency of a compression algorithm. Experimental results are plotted in terms of compression time and compression ratio for cipher texts of different sizes.

3.1 Compression Time

The algorithms implemented in Python. The execution time required to compress a cipher texts with different sizes are shown in Table 1.

From Fig. 1, it can be seen that LZMA compression method works faster even for a large file.

3.2 Compression Ratio

Compression ratio is one of the evaluation parameter of any compression algorithm. It is the ratio between the compressed file and the original file.

$$\text{Compression Ratio} = \frac{\text{Compressed File Size}}{\text{Original File Size}} \tag{1}$$

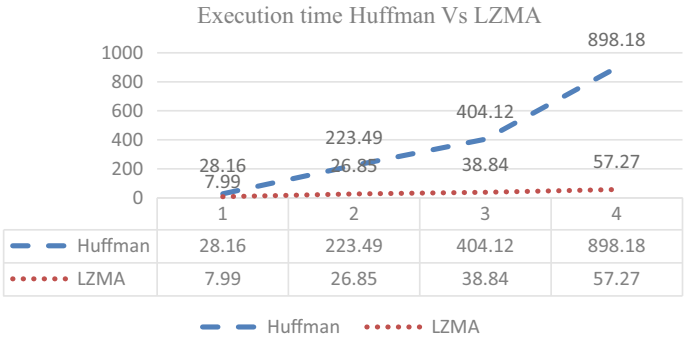


Fig. 1 Execution time Huffman versus LZMA

Huffman and LZMA compression methods were applied on a cipher texts of different size to determine the effect of compression on a cipher text. The tabulated results of compressed cipher text size are given in Table 2. The graph portraying the cipher text size before and after compression is shown in Fig. 2.

Table 2 Compressed cipher text size

S. No.	Cipher text length before compression (KBs)	Cipher text length after Huffman compression (KBs)	Cipher text length after LZMA compression (KBs)	Compression ratio (Huffman)	Compression ratio (LZMA)
1	4527	2422	1832	0.53	0.40
2	13,534	7221	5462	0.53	0.40
3	17,939	9577	7239	0.53	0.40
4	26,821	14,317	10,821	0.53	0.40

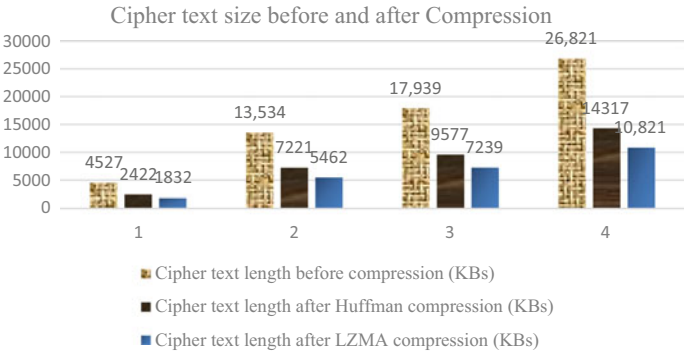


Fig. 2 Cipher text size before and after compression

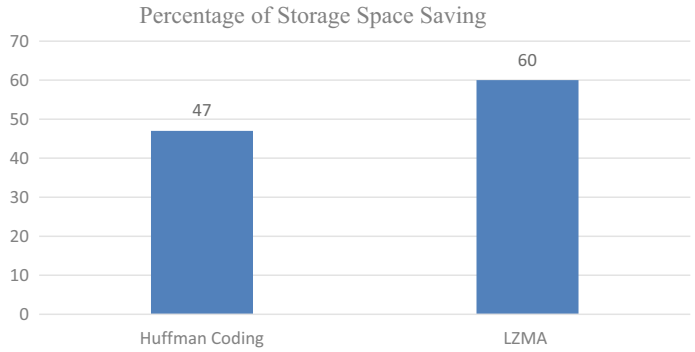


Fig. 3 Storage space savings Huffman versus LZMA

Storage space savings can be measured using the formula given below:

$$\text{Space Saving} = 1 - \frac{\text{Compressed File Size}}{\text{Original File Size}} \tag{2}$$

Figure 3 shows the storage space savings achieved using Huffman coding and LZMA are 47% and 60%, respectively.

4 Conclusion

In this work, the effect of compression on cipher text toward improved storage is studied. Cipher texts with different sizes are compressed using two major types of compression methods, i.e., Huffman algorithm based on entropy coding method and LZMA which is a dictionary-based compression algorithm. From the results, it is clear that cipher text can be compressed for storage proficiency. Huffman coding and LZMA algorithm resulted in a 47% and 60% of storage efficiency, respectively. Thus, this supports the idea that combination of encryption and compression provides a secured and optimized data storage.

References

1. Massoudi A, Lefebvre F, De Vleeschouwer C, Macq B, Quisquater J-J (2008) Overview on selective encryption of image and video: challenges and perspectives. Eurasip J Inf Secur 2008(1):179290
2. Finnila C (1994) Combining data compression and encryption. In: WESCON/94 "idea/microelectronics" conference Rec, Anaheim, CA, pp 404–409
3. Franceschini R, Mukherjee A (1996) Data compression using encrypted text. In: Proceedings of the third forum on research and technology advances in digital libraries, IEEE, pp 130–138

4. Bose R, Pathak S (2006) A novel compression and encryption scheme using variable model arithmetic coding and coupled chaotic system. IEEE Trans Circuits Syst I Regul Pap 53(4):848–857
5. Klein ST, Bookstein A, Deerwester S (1989) Storing text retrieval systems on CD-ROM: compression and encryption considerations. ACM Trans Inf Syst (TOIS) 7(3):230–245
6. Zhou J, Liu X, Au OC (2013) On the design of an efficient encryption-then-compression system. In: 2013 IEEE international conference on acoustics, speech and signal processing, IEEE, pp 2872–2876
7. Jiang J (1996) Pipeline algorithms of RSA data encryption and data compression. In: Proceedings of international conference on communication technology. ICCT'96, vol 2. IEEE, pp 1088–1091
8. Salomon D (2000) Data compression: the complete reference, 2nd edn. Springer, New York, Inc

RSS Using Social Networks User Profile

S. M. Manasa, Shreesha, Leena Giri, R. Tanuja, S. H. Manjula, and K. R. Venugopal

Abstract E-commerce application in recent days are gaining more importance due to its worldwide services and user-friendly applications. Some of the social networking applications are being backbone of e-commerce applications help in its functioning. With this, the number of e-commerce applications are being associated with social networking site and their user profiles. E-commerce sites use this social networking sites as service providers to their customers associated with their social networking profiles. In these online transactions, social networking sites act as third-party service providers to its users with respect to e-commerce interactions. Social networking sites acting as third-party service providers should be careful about users type, their behaviour and services that they provide. Users choosing their services should also be careful about what to choose and what not to choose. In this three-tier interactions, trust plays a very important role. Trust on any service providers should be evaluated to stay away from online fraudulent activities. User interacting through these applications should always be able to decide between the trusted services and fraud services. Hence, service selection in social networking is proposed with the help of *Refined Service Selection (RSS) algorithm* using social networks user profiles. Implementation is done with the help of real data set and found that the accuracy of *RSS* is more when compared with the existing algorithms.

Keywords Services · Social networks · Refined services · Social Profile

1 Introduction

Development in today's technology has gained so much popularity by its scalability on Internet applications while providing all the information required by users on their finger tips. Resources that are pooled to provide all kind of services to all the users worldwide. Resources can be any kind of information, links and services that

S. M. Manasa (✉) · Shreesha · L. Giri · R. Tanuja · S. H. Manjula · K. R. Venugopal
Department of Computer Science and Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bengaluru 560001, India
e-mail: smmanasa609@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_21

are available in digital form. This digital information is designed in such away that the information is served to all users through of applications that are associated with Internet. Users not only use the available information but they also try to provide the information required by the other users on different applications.

Social networks is one of the application that works associated with Internet. This applications have gained huge popularity and curiousity among the users. It is attracting users in millions to get registered on its applications on daily basis. Number of activities that are taking place on social networks are of different domain. Activities including social work, e-commerce, small-scale online business and user interaction. These activities involve users and their information shared on social networks. Hence, providing trust to these kind of activities on social networks is major challenge, as the complete user profiles of all users are not revealed due to security issues.

Service-oriented online social networks (SOSNs) are platforms that play an important role in providing various services including e-commerce. People in need of any services should be aware of choosing trustworthy service providers before availing any kind of services. Trust evaluation of service provider plays crucial role while availing the services. Service-oriented social networks (SOSNs) is association of social networks with e-commerce. Social networking group people together so that they can exchange their views, they can communicate and make new friends. Social networking is generally a structure of users related to each other. Some of the examples of social networking sites as Facebook, Twitter, Google plus and Instagram.

In Web services network, the structure of network mainly consists of individuals called participants or organisations, interactions between them are called links, relationship and connection. In this kind of network structures, trust is the main factor to be considered in successful maintenance and operations of SOSNs. SOSNs should have more trust factors to be satisfied when compared to traditional OSNs. This helps in proper maintenance of relationship between the participants that are more complex, and more specific in social networks. Traditional OSNs are not capable of evaluating the information based on common intelligence and human behaviour while SOSNs are more likely dependent on context aware techniques which includes Web analysis, social search and social recommendations. Single trust evaluation metrics used in OSN are not efficient in providing trust information between the participants. To evaluate this, SOSNs came to existence where Web services are merged with OSNs to evaluate trust between service provider and consumers. Choosing optimal social trust path (OSTP) becomes easy and yields too most trustworthy evaluation result.

1.1 Motivation

In real world, social networks are being used by many users on daily basis for information retrieval and entertainment. In recent times, online purchasing is the priority of many other people, and combining social networks with e-commerce is making user job easy; it is being helpful for many users on social networks. To

make these transactions secure and reliable, choosing a trustworthy service provider and their link analysis is challenge in social networks. To overcome the challenge of choosing trustworthy service provider, *Service Selection Algorithm (SSA)* is proposed to choose refined services that are trustworthy.

1.2 Contribution

To be aware of all the fradulent activities while choosing trustworthy service provider, *Refined Service Selection Algorithm (RSS)* is proposed. Social trust evaluation is carried out between service provider and consumer with the usage of direct relationship and recommendation. The products available in the Web services are rated and recommended mainly with the influence of service providers.

1.3 Organisation

The rest of the paper is organised as follows. Related work is differentiated in Sect. 2. Section 3 describes problem statement. Section 4 describes system overview SOSNs structure. Section 5 Matching User Social Profile. Section 7 result analysis is discussed. Paper conclusion is done in Sect. 8.

2 Related Work

The Web transactions have to make sure the online content that is being shared must be among the trusted candidates to achieve the propagation of ideas, content originality, service provider selection and for the reliability of the content posted on online social networks. To provide the above trust in transactions, Kim et al. [1] has designed a recommendation method that utilise the fast spreading and information sharing capability of a large customer network. Recommended method *customer-driven recommender system (CRS)* uses the *collaborative filtering (CF)* principle that computes a recommender list with the help distributed and local searches for same neighbours in a customer network. The experimental results are conducted on two data sets consisting of digital content of tangible products on showing that CRS is as accurate as global search mechanism of the traditional recommender system. The recommender list is computed rather it can be used to generate trusted list to the Web to make sure the system is trusted in nature.

Service-oriented social networks provide huge number of services to the users on social networks. The number of services are from different service providers and lot of paths are designed. It is a challenge to choose one optimal social trust path without any overlaps. To address this issue Liu et al. [2, 3] has presented the con-

cept of complex social network and *Quality of Trust (QoT)*. Then, a model design is built to select k optimal social trust paths ensuring end-to-end QOT called *Multiple constrained K optimal paths (MCOP-K)* selection problem, and with the help of Dijkstra's algorithm, a social optimal path is selected. The results obtained are qualitative in identifying the social trust paths compared to existing methods.

In social networks, it is the user responsibility to be careful about the other users before initiating any type of interactions, and it is a challenge to be safe online. As a solution, Li et al. [4] considers two factors, degree of centrality and contact interval of user to design trust evaluation model (T-OSN). T-OSN aims in proper evaluation of trust on online social networks that is reliable and more efficient. The advantage of this model is that it does not depend on traditional methods to evaluate the trust. Model designed is simple but effective to obtain the results using the model formulations. More the number of features, more will be trust score propagation.

Selecting a optimal social trust path that could actually help in evaluating trust of the service provider is the primary requirement in SOSNs. Existing single trust value evaluation methods provide good and basic information about the service providers. To know the social intimacy degree and role impact factor between the intermediate nodes, single-level trust value evaluation is not enough. To gain more information, Liu and Jia [5] propose *path integral Monte Carlo quantum annealing (PIMCQA)*-based *OSTP (PIMCQA-OSTP)* selection algorithm for complex SOSNs. *PIMCQA-OSTP* shows its accurate search ability and performs better than existing methods. Disadvantage is that the evaluating trust levels will include few more parameters to work with rather than considering only two parameters.

Path dependency makes it difficult to aggregate trust values for multiple dependent trust paths. As there are many paths existing between two same users, trust propagation decay can be seen in each node. These two challenges together known as generalised network flow problem. To overcome this challenge Jiang et al. [6] proposes a modified flow-based trust evaluation scheme called *GFTrust*. This deals with path dependencies using network flow and takes care of trust decay which means leakage that is associated with each node. Experiment is conducted on real social network data set of epinions and advogato. Results discussed can predict trust in online social networks with high accuracy. This social trust propagation path selection is a challenge in e-commerce site.

Social networking is such a platform where users tend to interact with strangers most of the times. Because of the biased behaviour of the strangers online, it is very important to access the trustworthiness of the user before initiating any type of conversations. This decision plays a very important role while influencing online. To assess trustworthiness of the user, Kianinejad and Kabiri [7, 8] propose a trust inference method that learns about trust value propagation among trust relationships to estimate the reliability of a stranger who want to trust him/her. Simple process to calculate the trust values using a new entity in data set is the advantage of this method. The experiments are conducted on real-time data set called Ciao, results explain that the accuracy in evaluating trust values is higher that TidalTrust, MoleTrust methods.

Authorisation for accessing secure information by a larger community of users in online social networks is challenge in today's Internet. Computational dynamic trust

model is designed by Zhong et al. [9, 10] for user authorisation in social networks. This model aims at assessing trustee's different trust values at different contexts from different trustors. Simulation results shows that the performance of proposed model with other trust models for different user behaviour patterns. Results show that the proposed model achieves higher performance than other models while predicting the behaviour of biased users.

3 Problem Statement

Problem with the existing system is that social networks are dependent upon single trust value evaluation. While this fails in terms of trust evaluation in complex OSNs. This leads to the services that are not trustworthy to be recommended by the system and results in security threat to the user.

4 System Overview

Service-oriented social networks is the combination of traditional social networks and Web of services or service providers. The system architecture consists of modules that work on parallel computing. Each module in the architecture performs its own functionalities and provide efficient output. Social networks consists of users in large number, and each user will have their profile created by themselves with all user details. Service providers will have all the links that give access to the services on any platform. Availing these services on online social networks has to have authenticated users and their registrations (Fig. 1).

Online social networks module is an application that has huge number of users using the application. This application allows any number of interested users to use the application. Each user can have any number of accounts on online social networks

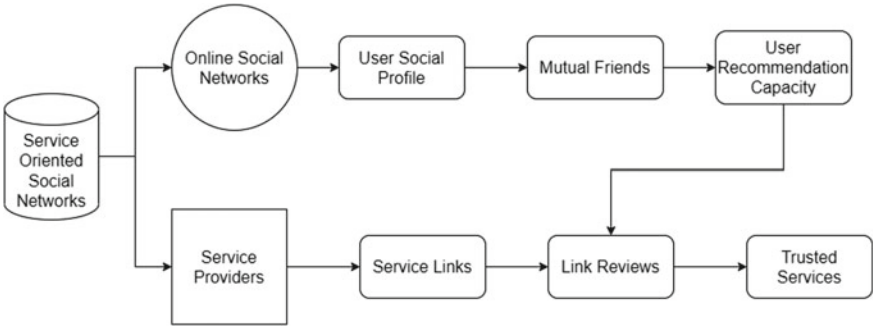


Fig. 1 System architecture for refined service selection in SOSNs

but should have different email id while registering on the application. The registered email id is the primary key while registering online. Same user with different mail ids can create different accounts on this application. User social profile module gathers the user information given on application during their registration. Each users basic information is accessed with the help of user social profile module.

Mutual friend module identifies the user with similar interest. Social network consists of wide variety of user with different interest. To identify the target users mutual friend module is designed. Mutual friend or target users are calculated with the help of cosine similarity between the users. Cosine similarity of the users determine the similar users using social networking application. To determine the mutual friends each user profile details, their interest and behaviour pattern on social networks are analysed. Each user profile, interest and behaviour pattern between the fellow users are analysed, similarity between the users are computed. Each user with similar user are termed are target users.

Service-oriented social networks send users with lots of links that provide different kind of services. These services can be anything that match user behaviour on social networks. Not all service links that are sent will be secure and trusted services; hence, these services should be verified, authenticated and approved by known users on the applications. Approval on the links or the service providers in social networks are done by reviewing the links are reviewing the services. Link reviews of each target user is analysed in link review module. Any link recommendations received on social networks, first step of authenticating is done by checking the target user review.

Once the target user reviews are verified and found to be good review, then that link and its intermediate node details are saved. Now, each target node and its intermediate nodes recommendation capacity are calculated. If the recommendation capacity of target node and the intermediate nodes are more than 0, then that service is considered as trusted service and notifies users to user the recommended link without compromising on the security issues in online social networks.

5 Matching User Social Profile

Social network is an interaction platform that depends on the user relation and their activities. It is difficult to analyse the similar users as huge number of users participate in social networks. To calculate the similarity between the users, profile attributes of each user is considered. Each user profile attributes will be a bag of activities. These activities are future documented for each user. These documentation is matched with other users on social networks. Cosine similarity Eq. 1 is used to calculate the similarity between the users on social networks. Each user document is a non-zero vector, and then, the similarity between these vectors are calculated. Number of similar users are termed as the target users.

Similarity metric $\cos\theta$ is the similarity of profiles attributes between the users given by

$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{||\vec{a}|| \cdot ||\vec{b}||} = \frac{\sum_1^n a_i b_i}{\sqrt{\sum_1^n {a_i}^2}\sqrt{\sum_1^n {b_i}^2}} \quad (1)$$

where a and b are the users on social networks. Every users in the network has a profile attributes that are considered as a individual vector for each user say P_a is the Profile attribute vector of user a, and P_b is the profile attribute vector of user b.

6 Service Selection Using Target Users

Choosing services of social networks with the help of mutual friends is done with the help of similarity of services algorithm. List of intermediate nodes and their information is calculated by Eq. 1. Each intermediate node's list of services are fetched from the user activity and their links are stored in the database. Generate the replicas of source user services with the series number. Source user service list replicas are distributed among the intermediate nodes. At any position of the randomly distributed replica check the number of neighbour nodes. For each neighbour node, calculate the number of services matching from the distributed replica.

As the number of matching replicas increases, then the number of neighbour gets mutually updated. For the dynamically updated neighbour nodes services, list has to be verified. For verifying the list of services, timespan of each service link is computed. Timespan of the link is the time difference that exists between the link recommended and link accessed. Any link sent on social networks can be accessed anytime by the user. This nature of the links can be sent to target user to get compromised on social networks. It is necessary to be careful while accessing the actual links that has shorter timespan. Timespan of the service link that is within a day is verified and added to the database and termed as verified. Then, the verified users trust value with the source user is set 1. Any recommended link matching with the verified service link in list is the mutual service a user can access on social networks.

7 Result Analysis

Real data set is used for the experimental results. The smaller data set is used for verifying the accuracy of the algorithm implemented. As social networking data is a huge number, it is complex for computation. Implementation of the work is done in Python. Similarity of the users and the number of services will lead to deploying more number of neighbour nodes which helps in finding out the reliable link in choosing the optimal social trust path. Reason to vary the number of neighbour nodes is that the number of paths that are generated can be analysed and the parameters to set the node dense can be improved. From Fig. 2, it is shown that the varying number of neighbour nodes gives the number of different social trust paths from the network.

Algorithm: Service selection algorithm

Input : List of intermediate nodes
Output: List of mutual services, link time span and list of recommender
begin
Initialization: set Time span = 0, List of link = 0 and Service list = $\{S\}$
Step 1: **while** T *!= 0* **do**
Randomly disturb the list of services
Step 2: **for** $r = 1 \ldots p$, **do**
Select the replica in position h
Step 3: **for** $N_i = 1 \ldots M * N - 1$ **do**
Calculate the number of matching replicas with the existing neighbour
if $r = 0$ **then**
Total number of matching replicas $(R) = 0$
else
Total number of matching replicas $(R) = R + 1$
Step 4: Number of neighbor nodes are dynamically updated are equal to R.
Step 5: **for** M users $S1, S2 \ldots Sn$
$S_n(L_tn)$ = Link accessed time − Link recommended time
if $S_n(L_tn) <= 24$ **then**
Add the link to database as verified link
else
Discard the link
End

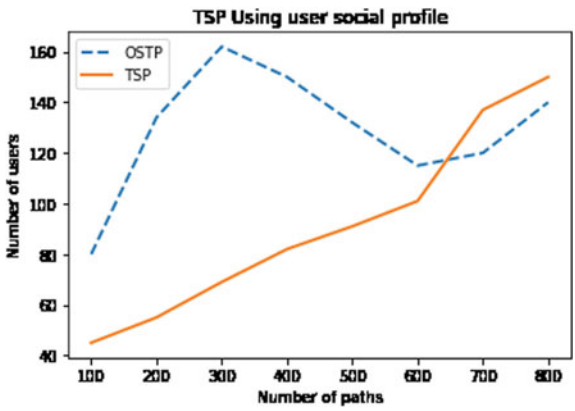


Fig. 2 TSP using social profile

8 Conclusions

In this paper, we have presented a complex social network structure that takes social network user and their reviews as the main attribute in computing the algorithm, reflecting the real-world situations better. In complex social networks, trust is the

major factor on which all users depend for accessing secure services. To analyse the better real-world situations on social networks, its relationships and recommendations, we have implemented the identification of services through mutual friends. Similar users are treated as mutual friends, this recommendations will make sure of recommending any service of interested domain. So that the recommended links are not spammed. Then, the path to select this service are based on the recommendation capacity of the node. Capacity of the each node ensure that no lame recommendations made inside the network. By this, the complexity of path formation is reduced.

References

1. Kim HK, Ryu YU, Cho Y, Kim JK (2011) Customer-driven content recommendation over a network of customers. IEEE Trans Syst Man Cybern Part A Syst Hum 42(1):48–56
2. Liu G, Wang Y, Orgun MA, Lim E-P (2011) Finding the optimal social trust path for the selection of trustworthy service providers in complex social networks. IEEE Trans Serv Comput 6(2):152–167
3. Gong NZ, Wang D (2014) On the security of trustee-based social authentications. IEEE Trans Inf Forensics Secur 9(8):1251–1263
4. Hao F, Li S, Min G, Kim H-C, Yau SS, Yang LT (2015) An efficient approach to generating location-sensitive recommendations in ad-hoc social network environments. IEEE Trans Serv Comput 8(3):520–533
5. Liu L, Jia H (2015) Trust evaluation via large-scale complex service-oriented online social networks. IEEE Trans Syst Man Cybern Syst 45(11):1402–1412
6. Jiang W, Wu J, Li F, Wang G, Zheng H (2015) Trust evaluation in online social networks using generalized network flow. IEEE Trans Comput 65(3):952–963
7. Kianinejad M, Kabiri P (2015) A strategy for trust propagation along the more trusted paths. In: 3rd conference on swarm intelligence and evolutionary computation (CSIEC), pp 1–6
8. Zhao G, Lei X, Qian X, Mei T (2018) Exploring users' internal influence from reviews for social recommendation. IEEE Trans Multimed 21(3):771–781
9. Zhong Y, Bhargava B, Lu Y, Angin P (2014) A computational dynamic trust model for user authorization. IEEE Trans Dependable Secure Comput 12(1):1–15
10. Hampi KSH, Rao NPC, Kumar SP (2015) A computational dynamic trust model for user authorization

Learning-Based Framework for Resource Allocation for Varying Traffic

P. Anvita, K. Chaitra, B. Lakshmi, Shraddha, B. Suneeta, and P. Vishal

Abstract Network function virtualization (NFV) presents a model to remove physical middle-boxes and replace them with virtual network functions (VNFs) that are more resilient. Adjusting resource allocation in response to the varying demands of traffic, there is a need for instantiating the VNFs and also balancing the resource allocation based on demand. Current optimization methods frequently expect the amount of resources required by every VNFs instance is fixed, resulting in either resource wastage or poor quality of service. To resolve this issue, machine learning (ML) models are used on real-time data of VNF, which contains performance indicators and requirement of resources. Evaluating the result, using ML models along with the VNF placement algorithms shows the reduced amount of resource consumption, thereby improving the quality of service and reducing the delay.

Keywords Resource allocation · Machine learning · VNF · QoS

1 Introduction

Determining the quantity of resources to be allocated is a difficult task. Allocating a fixed number of resources can easily result in over or under-allocation. Existing resource allocation techniques presume that each VNF instance has fixed resource requirements; however, if less number of resources are allocated, it leads to bad service quality, and if more number of resources are allocated, it leads to wastage of resources. Many cases of excessive heavy traffic, very high connection density and mobility pose challenges to mobile networks. The goal of resource allocation problem is to determine the best way to allocate fixed amount of resources to activities while minimizing the cost.

Resource allocation supports to guarantee that the quality of service standards is satisfied. It agrees to ensure a high standard of security by contradicting different

P. Anvita · K. Chaitra · B. Lakshmi · Shraddha · B. Suneeta (✉) · P. Vishal
KLE Technological University, Hubli, India
e-mail: suneeta_vb@kletech.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_22

denial-of-service attacks. By allocating fixed amount of resources which results in under-allocation and over-allocation of resources. Network function virtualization: In NFV, network services contains many VNFs, at available points of presence (PoPs), they are positioned. Because of many aspects of the network service, the underlying substrate network, as well as demand of current service, dynamically positioning these network services and VNFs is critical. The VNF algorithm needs to focus on amount of resources to assign to every instance along with number of VNF instances and location of it. Every VNF is conservatively instantiated with a pre-defined number of resources specified in the VNF descriptor by the VNF developer. VNFs, on the other hand, do not require continuous resources and instead require too many or fewer resources based on traffic load. As a result, allocation of set volume of resources might result in resource under or over-allocation. When it comes to under-allocation, VNFs are deficient of resources to process all packets, which causes packet drops and lowers service quality, if not entirely destroys service functioning. As a result, by allocating large amount of resources under-allocation of resources can be avoided. Again as a result, over-allocation occurs, squandering resources, preventing other VNFs to use it, and resulting in unnecessarily higher costs.

2 Related Work

In paper [1], a network slicing framework which consists of resource management system is proposed and implemented over all the network slices over 5G. The major obstacles occurred during network slicing is also discussed. The network slicing framework consists of three layers and can easily deployed and handled different network applications over one common infrastructure. The paper combines different ML algorithms along with NFV and SDN, so as to create an experimental framework and architecture for the upcoming self-organizing networks and network slicing techniques as proposed in paper [2]. A software-defined network (SDN) controller ensures that every application is preserved with a relevant bandwidth. A clustering model is used for grouping similar applications together. This clustering model is used to build a classification model so that the incoming traffic loads can be classified and sent to appropriate slices with the help of SDN controller.

In paper [3], the model DEEPSLICE which is derived from deep learning techniques is implemented to handle traffic loads and availability of the network. The model's major goals are to choose the right network slice for a machine, make accurate slice predictions, then allocating enough resources to the specific slice based on traffic predictions. Paper [4] discusses the basic concepts of deep reinforcement learning (DRL), and its use in resource management for network slices. This shows the advantage of DRL in managing resource allocation in different slicing cases. The authors have discussed about yet another common problem, slicing of the main network depending on priority for common VNFs. Some important problems mentioned are: slice admission control; hidden actions and states; latency and accuracy to retrieve rewards; policy learning cost. Disadvantages: lacks necessary learning speed.

Authors of the paper [1] have described the design problems of the network slicing in 3GPP, and NextGen RRC architecture, as well as a new network slicing architecture have recently been discussed as a solution to solve the existing issues. Paper [5] presents a designs of ML for allocation of resources and way to find patterns hidden in a vast amount of data. The optimal or near-optimal solution of a nearer historical case is utilized for allocation of resources for the current scenario using the extracted similarities. Paper [6] discusses the way ML can be used to derive models and techniques straight from real VNF data that can determine requirements of VNF resource exactly based on traffic load. The use of models in VNF placement algorithms will have an impact on the quality of the VNF placements that exist. By comparing ML algorithms and a general fixed resource allocation approach, shows selecting a suitable approach for dynamic resource allocation is important.

3 Proposed Framework with Learning-Based Algorithms

The details about the learning framework used is presented. Before demonstrating chosen algorithm, we have carried out a pre-requisite work, which is to assess ML methods for resource allocation. As mentioned assessment consists of using Squid and Nginx dataset's performance measurements. These contain millions of data points and also a subset of various resource configurations. To assess the six different ML models—linear regression model, decision trees, SVR, ensemble learning, and neural networks. They are compared with each other across the various resource configurations. Also, these models are compared with another model that is the common model which assigns a fixed predetermined amount of resources to the VNF instances. Before the training step, preprocessing of the data is done. The data is filtered into a subset of required features only (measurement of maximum sustainable traffic load for the various CPU configurations). The features not relating to resource allocation are assigned a fixed value. This process of feature scaling is important for the ML models. After preprocessing, the models are trained and evaluated using the fivefold cross-validation.

3.1 Impact on the Allocation of Total Resources

Figure 1 depicts the overall amount of allocated resource as the traffic load increases for these three algorithms used. The VNF instance which are placed, require more amount of resources in order to manage the boosting traffic. The total number of allocated resources will depend on the requirement of prediction of resources, that is *rc*, for every model. It demonstrates that when a pre-defined quantity of CPU resources is assigned to all the VNF instances, will rapidly fill up the capacity of network's resources. That is, fixed resources will lead up to 12 times increase in the allocated number of resources of CPU in SVR and gradient boosting denoting

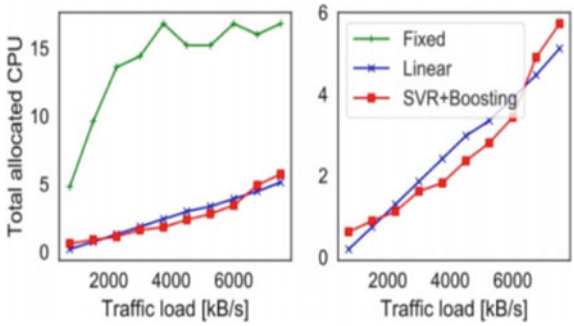


Fig. 1 Total allocated CPU in VNF placements computed with different models

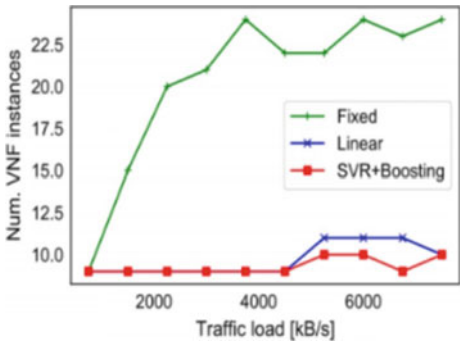


Fig. 2 Model impact on the number of VNF instances

resources allocated in excess. Whereas, machine learning models, combining linear approximation and SVR show less problems.

Effect on Number of VNF Instances: As the resource allocation is speeding, more instances of VNF must maintain the load across the various nodes. Figure 2 shows that the said statement is true with respect to fixed resource allocation, where almost twice the number of instances are placed compared to SVR model and the gradient boosting technique.

3.2 ML Assisted with Cloud Computing: Mathematical Modeling

The background for wireless communication systems that are aided by cloud computing comprises of the following three elements:

– A cloud system: consisting of configurable computing resources which are assembled into as cloud with vast storage potential and excellent computational services.
– Base station: which provides wireless access functions.

– Backhaul links: the estimated data is fetched from the base stations to the cloud administration by these backhaul links. They also organize the system of resource allocation at the base stations.

The resource allocation which is occurring at the base station can be expressed in the form of a mathematical expression as shown below:

$$\text{minimize } g(y, b);\ y \in S \tag{1}$$

$$\begin{aligned} \text{subject to } & h_i(y, b) \leq 0; \quad i = 1, \ldots, m \\ & j_i(y, b) = 0; \quad i = 1, \ldots, p \end{aligned} \tag{2}$$

In the equation, y represents the vector of variables of the optimization problem, the objective function which is to be minimized by the vector y is $g(\,)$, the vector that describes the problem is b, ${h_i}^m{}_{i-1}$ and ${j_i}^m{}_{i-1}$ are called the inequality and equality limited functions, respectively, and S is called a constraint set. It is evident that the scenario is a minimization problem. A maximization problem can be assumed by negating the objective function of the optimization problem. By viewing the resource allocation problem in the form of an equation as mentioned, all of the values in the vector y are denoted as variables that define amount of allocated resources, which include the transmit power and the sub-carrier index. All of the values in the vector b are attributes of the system or parameters of the wireless broadcast, such as bandwidth, background noise level, and the sub-carrier number. ${h_i}^m{}_{i-1}$ and ${j_i}^m{}_{i-1}$ are used to define the detailed case and the restrictions on resource allocation, which are the users' QoS demands, effect from all the possible noise and interference, and accessibility of resources. The objective function represents the features of the finest solution available and defines the key objective which is the metrics for execution of resource allocation. The specific case that is described by b, the ideal solution for resource allocation y^* is that vector which gets a perfect value of objective function midst the remaining vectors and surpasses all the limitations.

3.3 Machine Learning Framework

Wireless communication systems which are aided by cloud computing require a vast quantity of data on historical circumstances, which are gathered together and deposited in the cloud system. The idea behind resource allocation system using ML algorithms is classifying these solutions and extracting the masked comparisons in these historical situations. This resource allocation model will be forwarded to analyze a base station and allocate resources more efficiently. Figure 3 shows the schematic representation of the proposed ML framework. A massive quantity of historical data on situations is stocked in the cloud using the cloud storage system. This data contains a large number of attributes, i.e., the user number, the channel

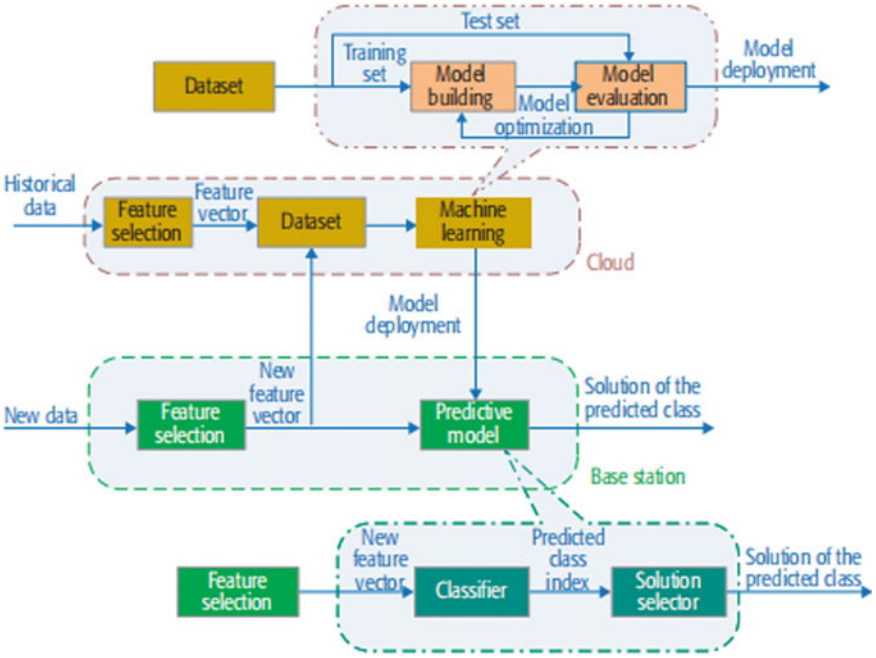


Fig. 3 The designed framework for resource allocation using ML framework

state information (CSI) of users, etc. Further on, feature selection is performed on the data to distinguish and remove unimportant and irrelevant features from the data. Also, the process of feature selection selects certain features or parameters from the historical data and presents them in the form of a feature vector. These feature vectors are composed together to establish a huge dataset of required features. Lastly, a prediction model is built with this data which makes a decision for allocating the resources for any future situations.

3.4 Feature Selection

Feature selection returns a subclass of the existing features, or variables, or predictors. The selected features contribute the maximum to the prediction output of the network or model. This process can increase the data in the training set as the dimension of each feature is added into a feature vector. Therefore, it is imperative to obtain a small dimension of the feature vectors that accurately capture the principle of resource allocation in our required scenarios. To do so, only the important and required information for resource allocation should be selected as the subclass of features. As we can see in Eq. (1), resource allocation is demonstrated as a optimization problem and all the required and important information are contained within vector a.

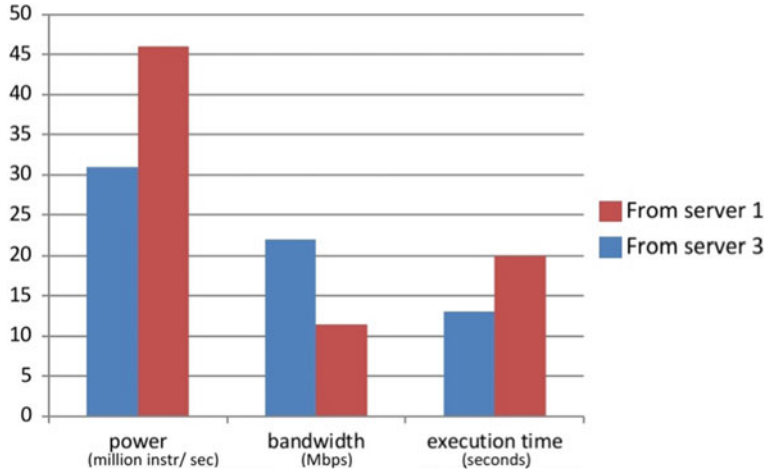


Fig. 4 Allocated resources for user 1

4 Results

The cloud module is responsible for storing raw data, preprocessing the data, managing the warehouses, learn based on the data, and create ML model based on it and finally sending model to edges. Data is stored in its storage system, and here, we do the whole learning process in the cloud.py module. The module reads the data, splits it on 70% for training and 30% testing, trains its model using decision tree classifier technique, tests its model and persists its model in the storage system. The classifier technique is chosen using cross-validation technique which showed that decision tree can classify the data with 85% accuracy which is a good number for resource allocation. Other techniques like KNN had about 65% accuracy on out data. Tasks are generated on different areas (in our application we have three different position). Tasks are generated on different areas (in our application we have three different position). Worker is our resources, machines which do the jobs. Our application has three workers and each worker has some attributes:

- Position: Position of the worker server
- Power: Power of the worker server in million instructions per seconds
- Bandwidth: Bandwidth between the worker server and other workers in mbps
- Makespan (timer): The time the worker consumed to do the jobs in second.

Figure 4 shows the allocated resources for user 1. The given resources are power, bandwidth, and time required for execution. Figure 5 shows the allocated resources for user 2. Figure 6 shows the allocated resources for user 3. Figure 7 shows a comparison graph between our proposed ML algorithm and greedy algorithm. As we can see from the graph, ML algorithm uniformly distributes the resources for all the three users, whereas greedy algorithm shows non-uniformity across users.

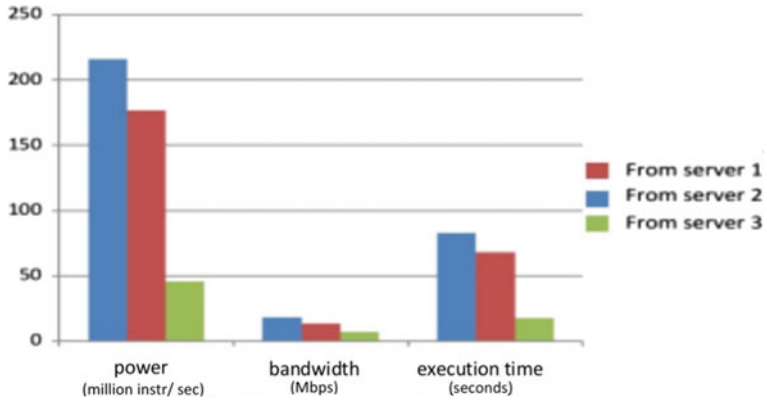


Fig. 5 Allocated resources for user 2

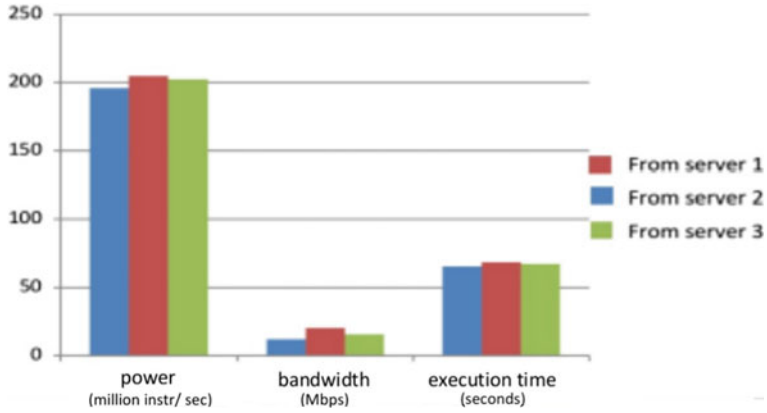


Fig. 6 Allocated resources for user 3

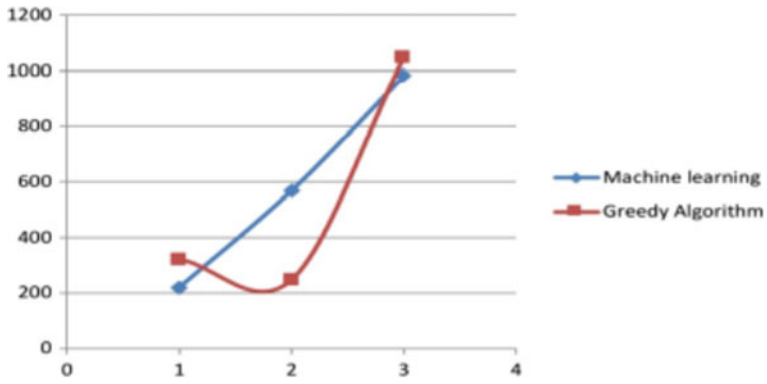


Fig. 7 Comparison graph between our proposed ML algorithm and greedy algorithm

5 Conclusions

Resource allocation by traditional techniques face considerable difficulties in the near future to address the ever-increasing QoS demands of users with limited radio bandwidth. Therefore, dynamic resource allocation using ML reduces the use of resources and improve quality of user service. As a result, including ML models into placement algorithms for VNF will have a considerable effect on the standard of the final placement of VNF. We introduced a ML model for radio resource allocation and discusses the application of supervised learning algorithm to determine features with similar patterns present in data of many of the past scenarios, and for the radio, resource allocation in the current scenario is used.

References

1. Yoo T (2016) Network slicing architecture for 5G network. In: International conference on information and communication technology convergence, pp 1010–1014
2. Le L, Lin BP, Tung L, Sinh D (2018) SDN/NFV, machine learning, and big data driven network slicing for 5G. In: IEEE 5G world forum, pp 20–25
3. Gumaste S et al (2020) Detection of DDoS attacks in OpenStack-based private cloud using Apache Spark. J Telecommun Inf Technol 4:62–71
4. Li R et al (2018) Deep reinforcement learning for resource management in network slicing. IEEE Access 6:74429–74441
5. Schneider S et al (2020) ML for dynamic resource allocation in network function virtualization. In: IEEE conference on network softwarization, pp 122–130
6. Budihal SV et al (2016) Framework for intercell interference avoidance in MIMO OFDMA system. In: International conference on wireless communications, signal processing and networking, pp 1626–1630
7. Budihal S et al (2018) Framework for dynamic resource allocation to avoid intercell interference by evidence theory. In: Advances in intelligent systems and computing, vol 518, pp 391–402
8. Budihal SV, Banakar RM (2020) Evidence-based dynamic radio resource allocation to mitigate inter cell interference employing cooperative communication. IET Commun 14(12):1848–1857

Machine Learning Classifiers for Detecting Credit Card Fraudulent Transactions

Bharti Chugh and Nitin Malik

Abstract Credit card usage has increased significantly as a result of the fast development of e-commerce and the Internet. As a consequence of enhanced credit card usage, credit card theft has risen substantially in recent years. Fraud in the financial sector is expected to have far-reaching effects in the near future. As a response, numerous scholars are concerned with financial fraud detection and prevention. In order to prevent bothering innocent consumers while detecting fraud, accuracy has become critical. We used hyperparameter optimization to see if created models utilizing different machine learning approaches are significantly the same or different, and if resampling strategies improve the suggested models' performance. The hyperparameter is optimized using GridSearchCV techniques. To test the hypotheses of data that has been divided into training and test data, the GridSearchCV and random search methods are used. The maximum accuracy 72.1% was achieved by decision tree classifier on the imbalanced German credit card dataset. The maximum accuracy of 98.6% is achieved by LDA on imbalanced European credit card dataset. Additionally, logistic regression and naïve Bayes were also tested and SMOTE was applied.

Keywords Decision tree · LDA · Gaussian Naïve Bayes · Logistic regression · Bernoulli Naïve Bayes · Credit card · GridSearchCV

1 Introduction

Major financial institutions are making their services available to the general public through online banking, mobile banking, credit and debit cards. Using services like credit cards, which have proven to be extremely good way for online purchases, makes everyday life easier. In the banking industry, credit card and online net banking fraud is a global issue. The credit card or any other card for the matter has data stored in a machine-readable format on a black stripe on the back. It includes details like the

B. Chugh (✉) · N. Malik
The NorthCap University, Gurugram, India
e-mail: bharti.kathpalia@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_23

cardholder's name, card number, expiration date, CVV code, card type, and other information that might be used to conduct credit card fraud. In the discipline of fraud detection with classifiers, financial fraud is a major issue. The assumption of balanced distribution in the dataset [1, 2] is a challenge for virtually all classifier learning methods. Machine learning techniques are used to anticipate various bank transactions. This article investigates the effectiveness of machine learning classifiers: logistic regression (LR), linear discriminant analysis (LDA), naïve Bayes (NB), and decision tree (DT). The algorithm's performance is evaluated using the recall score, accuracy, f1-score, and precision.

1.1 Literature Review

The default rate on credit loans across all commercial banks has been at an all time since last six years, according to Federal Reserve economic statistics, and it is expected to continue to rise into 2021. Duman et al. [3] sought to demonstrate the benefits of using data mining methods such as DT and support vector machines (SVM) to solve the credit card fraud recognition problem and reduce the banks' risk. The findings revealed that classifiers and added DT classifiers outperformed SVM approaches in tackling the problematic at hand. Wang et al. [4] presented a strategy for detecting credit card fraud based on local isolation coefficient to mining distance-based outliers on conventional algorithms. Bhattacharya et al. [5] detailed a comparative research on data mining methodologies but with the limitation of non-availability of exact time stamp data beyond the date of credit card transactions. APATE [6] is a new technique for detecting fraudulent credit card transactions in highly nonlinear models. A behavior-based credit card fraud detection model was proposed by Zhang et al. [7]. Chuang et al. [8] created a data mining-based model. Web services were utilized to communicate a fraud design, data communication between banks used to detect fraud. To identify credit card thefts, Yu et al. [9] suggested an outlier mining technique. Dembrani et al. [10, 11] proposed a comparative analysis of various adaptive filter structures that can be executed for credit card fraud recognition. A fusion method was presented [12, 13]. The four components were a rule-based filter, a Dempster-Shafer Adder, a transaction history database, and a Bayesian learner. Srivastava et al. [14–16] developed a hidden Markov model for detecting credit card fraud. They developed a unique credit card fraud detection system that uses best matching algorithms to detect 4 distinct patterns of fraud cases and addresses the associated difficulties reported by previous credit card fraud detection studies [17–19].

1.2 Organization of the Paper

The machine learning techniques applied to the proposed model are explained in Sect. 2. Section 3 depicts the suggested model's block diagram, flowchart, and entire implementation. Section 4 illustrates the comparative study with existing machine learning approaches. Section 5 discusses the conclusion and future scope.

2 Machine Learning Algorithms

2.1 Logistic Regression

The logistic regression model calculates a weighted sum of input characteristics and bias. Logistic regression is named for the function used at the core of the method, the logistic function. Any integer with a real value can be translated to a value between 0 and 1.The output value being modeled is a binary value (0 or 1) rather than a numeric number, which is a major distinction from linear regression. The logistic regression equation is shown below:

$$y = \frac{e^{b0+b1*x}}{1 + e^{b0+b1*x}} \tag{1}$$

where y is the expected output, $b0$ represents the bias or intercept term, and $b1$ represents the coefficient for a single input value (x). Each column in your input data has a b coefficient (a constant real number) that must be determined using your training data.

2.2 Decision Tree

Decision tree (DT) is a non-parametric supervised learning approach used for classification and regression. The objective is to learn basic decision rules from data characteristics to construct a model that predicts the class of a target variable. Instances are classified using decision trees by sorting them along the tree from the root to a leaf node, which yields the classification. Starting at the root node of the tree, an instance is categorized by testing the attribute given by its node, then proceeding along the tree branch according to the attribute's value. The sub-tree rooted at the new node is then processed in the same way.

2.3 Naïve Bayes

The naïve Bayes algorithm utilizes the Bayes theorem to classify the data. The naïve Bayes method essentially tells us the likelihood of a record belonging to a definite class constructed on the standards of its characteristics. Gaussian NB is a form of naïve Bayes that handles continuous data and follows the Gaussian normal distribution.

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(x_i - \mu_y\right)^2}{2\sigma_y^2}\right) \tag{2}$$

The parameters σ_y and μ_y are estimated using maximum likelihood.

The Bernoulli NB decision rule is based on the Bernoulli naïve Bayes decision rule:

$$P(x_i|y) = P(i|y)x_i + (1 - P(i|y))(1 - x_i) \tag{3}$$

2.4 Linear Discriminant Analysis (LDA)

The LDA model assumes that the data is normally distributed and estimates the mean and variance for each class. It is common to assume about this in the multivariate (single input parameter) case with two classes. Overall mean (μ) number of each input (x) for each class (k) may be found by dividing the sum of values by the total number of values.

$$k = \frac{1}{nk}(\text{sum}(x)) \tag{4}$$

The numeral of events with class k is nk, and the mean value of x for class k is μk. The variance (σ^2) is calculated to average squared modification of all value from the mean:

$$\sigma^2 = \frac{1}{(n-k)}\text{sum}\left((x-)^2\right) \tag{5}$$

3 Implementation

Figure 1 displays the block diagram of the proposed model. The suggested model's operation is sequenced as follows: data collection, data processing, research into the

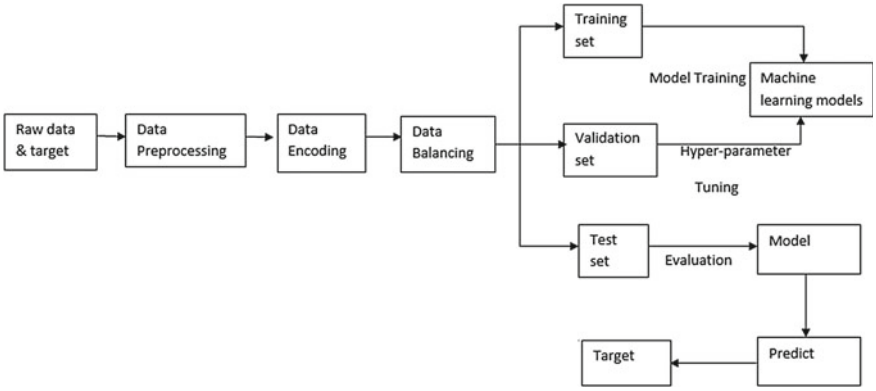


Fig. 1 Block diagram of the proposed model

appropriate model for the type of data, the model training and testing and evaluation. It is the most crucial stage in improving the accuracy of machine learning models. In supervised learning, an AI system is provided with data that has been labeled, meaning that each piece of information has been assigned to an appropriate label. Some of the most often used classification algorithms are support vector machine, naïve Bayes, logistic regression, decision trees, and KNN.

The dataset is divided into three categories as training data, validation data, and test data. To train the classifier, start with the training dataset, then fine-tune the parameters with the validation dataset, and lastly, evaluate the classifier's performance with the test dataset. The classifier has access only to the training and/or validation sets. The test dataset must not be used through classifier training. A testing set would be given mostly during the classifier's evaluation. Validation datasets are used to fine-tune the parameters of a classifier. When the categorical feature is ordinal, the categorical data encoding approach is used. The dataset is initially imbalanced, therefore, data rebalancing technique has been applied.

Figure 2 depicts the flowchart of the entire model stages followed by this technique. We have used two datasets: European and German datasets. The data is preprocessed and separated into two subsets: training and testing. The model is optimized using hyperparameter tuning, and the parameters such as accuracy, precision, recall, and F1-score are calculated using it.

4 Results and Discussions

The European credit card dataset [14] on which PCA technique has already been applied, contained 28 numerical features. The German credit card dataset [15] contained 21 features out of which 12 are categorical and 7 are numerical. The

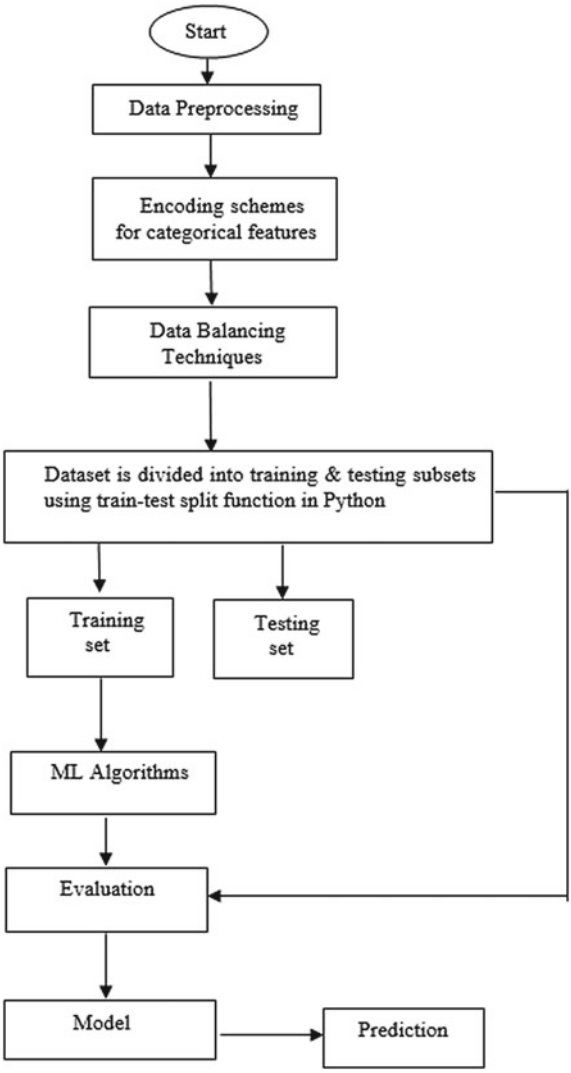


Fig. 2 Flow chart of the proposed model

proposed methodology is implemented in Python and uses machine learning classification methods. Several machine learning models such as LR, LDA, naïve Bayes, BernoulliNB, and decision tree are used to analyze it. Hyperparameter optimization is carried out using GridSearchCV (Table 3 and 4).

The results from the proposed methodology implemented using decision tree classifier are compared with that in Patil et al. [16] on German dataset which signifies the superiority of the method proposed (Table 7).

Table 1 Performance analysis without hyperparameter optimization on German dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 score
LR	71.50	69	71	0.69
LDA	69.50	66	69	0.66
Gaussian NB	71	73	72	0.73
Bernoulli NB	71.50	70	71	0.71
Decision tree	66.00	60	66	0.59

Table 2 Performance analysis without hyperparameter optimization on European dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 score
LR	99.80	100	100	1.00
LDA	98.50	97	95	0.97
Gaussian NB	91	92	93	0.92
Bernoulli NB	91.50	90	91	0.91
Decision tree	86	80	86	0.79

Table 3 Performance analysis using hyperparameter optimization on German dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 score
LR	69.50	64	69	0.64
LDA	70.7	66	69	0.66
Gaussian NB	71	73	72	0.73
Bernoulli NB	71.50	70	71	0.71
Decision tree	72.10	45	67	0.54

Table 4 Performance analysis using hyperparameter optimization on European dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 score
LR	98.5	100	100	1.00
LDA	99.6	95	97	0.95
Gaussian NB	91	93	92	0.93
Bernoulli NB	91.50	90	91	0.91
Decision tree	82.1	55	57	0.64

Table 5 Confusion matrix of decision tree classifier on German dataset

	Predicted No	Predicted Yes
Actual No	0	98
Actual Yes	0	202

Table 6 Confusion Matrix of LDA classifier on European dataset

	Predicted No	Predicted Yes
Actual No	56,852	9
Actual Yes	20	81

Table 7 Comparative analysis with the existing results

Decision tree classifier	Ref. [16]	Proposed
Accuracy	72%	72.10%

5 Conclusion

This article assesses the performance of different machine learning classification algorithms by means of a German credit card dataset to perceive whether or not an operation/transaction is fraud. The credit card dataset was imported, preprocessed, encoded, and equipped for training the model using the machine learning workflow mechanism. The models were verified using both hyperparameter optimization and non-hyper-parameter optimization methods. It was then trained, deployed, and assessed for each classification model using multiple parameters and assumptions. The decision tree classifier outperforms the LR, LDA, and naïve Bayes algorithms in terms of performance. Ensembling all utilized models utilizing voting ensemble or weighted average ensemble can help increase the model's accuracy.

References

1. Usama F, Gregory PS, Padhraic S (1996) From data mining to knowledge discovery in databases. AI Mag 17:37–54
2. Yanmin S, Wong AKC, Mohamed SK (2006) Classification of imbalanced data: a review. Int J Pattern Recognit Artif Intell 23(4):687–719
3. Şahin YG, Duman E (2011) Detecting credit card fraud by decision trees and support vector machines. In: International multiconference of engineers and computer scientists
4. Yu B, Song M, Wang L (2009) Local isolation coefficient-based outlier mining algorithm. In: International conference on information technology and computer science, vol 2. IEEE
5. Bhattacharyya S, Jha S, Tharakunnel K, Westland JC (2011) Data mining for credit card fraud: a comparative study. Decis Support Syst 50(3):602–613
6. Van Vlasselaer V, Bravo C, Caelen O, Eliassi-Rad T, Akoglu L, Snoeck M, Baesens B (2015) APATE: a novel approach for automated credit card transaction fraud detection using network-based extensions. Decis Support Syst 75:38–48
7. Yu W-F, Wang N (2009) Research on credit card fraud detection model based on distance sum. In: International joint conference on artificial intelligence. IEEE
8. Dembrani MB, Khanchandani KB, Zurani A (2017) Comparative analysis of various adaptive filter structures using Simulink. In: Proceeding of international conference on intelligent communication, control and devices. Springer, Singapore
9. Dembrani MB, Khanchandani KB, Zurani A (2017) FPGA implementation of low power & high speed accurate ECG signal detection using DABLMS algorithm. Communication and computing systems. In: Proceedings of the international conference on communication and computing systems (ICCCS 2016), Gurgaon, India
10. Guo T, Li G-Y (2007) Neural data mining for credit card fraud detection. In: International conference on machine learning and cybernetics, vol 7. IEEE
11. Panigrahi S et al (2009) Credit card fraud detection: a fusion approach using Dempster–Shafer theory and Bayesian learning. Inf Fusion 10(4):354–363
12. Srivastava A et al (2008) Credit card fraud detection using hidden Markov model. IEEE Trans Depend Secur Comput 5(1):37–48

13. Quah JTS, Sriganesh M (2008) Real-time credit card fraud detection using computational intelligence. Expert Syst Appl 35(4):1721–1732
14. Dal Pozzolo A, Caelen O, Johnson RA, Bontempi G (2015) Calibrating probability with undersampling for unbalanced classification. In: Symposium on computational intelligence and data mining (CIDM). IEEE
15. Statlog German credit card dataset, UCI Machine Learning Repository, https://archive.ics.uci.edu/ml/datasets/statlog+(german+credit+data. last accessed 2021/01/5
16. Patil S, Nemade V, Soni PK (2018) Predictive modelling for credit card fraud detection using data analytics. Procedia Comput Sci 132:385–395
17. Itoo F, Singh S (2021) Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. Int J Inf Technol 13(4):1503–1511
18. Thennakoon A, Bhagyani C, Premadasa S, Mihiranga S, Kuruwitaarachchi N (2019) Real-time credit card fraud detection using machine learning. In: 9th International conference on cloud computing, data science & engineering (confluence), pp 488–493
19. Cynthia PC, Thomas George S (2021) An outlier detection approach on credit card fraud detection using machine learning: a comparative analysis on supervised and unsupervised learning. In: Intelligence in big data technologies beyond the hype. Springer, Singapore, pp 125–135

Virtual Ornament Room Using Haar Cascade Algorithm During Pandemic

S. S. Sunaina M, L. S. N. J. Manjusha P, and Kishore O

Abstract Generally, purchaser wishes to attempt all jewels on them and share those photos on social media to get a recommendation. But now, in the present situation, all are avoiding touching activities unnecessarily. It is not so safe due to COVID-19. So without stopping purchaser jewellery attempts, a solution is provided by using an application called virtual ornament room. The proposed scheme is based on creating a Web application that identifies where the human face is located in the frame and superimposes the chosen jewellery on the face using the HAAR Cascade Algorithm. By using the concepts of augmented reality, the digital objects are estimated and placed onto the frame in real time. The system is implemented by using Flask framework and OpenCV, a Python module. This application works with an attached camera, Internet, and a Web browser. From the results, it can be seen all the jewellery can be easily selected by the customers from their home itself during the pandemic situation.

Keywords Augmented reality · Virtual ornament room · Flask framework · HAAR cascade

1 Introduction

Jewellery has played a crucial role in human life from thousands of years. Many women like to wear jewellery as a symbol of femininity. Jewellery can also make a woman feel confident and beautiful. Not only for women, jewellery indicates an individual status for men. Everyone wants to try the jewellery before the purchase, and this can be done by the proposed application [1]. This application helps the customers to try on the piece of jewellery virtually without any physical contact. Using, augmented reality technology with the help of the HAAR cascade algorithm, where a superimposed digital ornaments are projected on their face in real time. The

S. S. Sunaina M (✉) · L. S. N. J. Manjusha P · Kishore O
Aditya Engineering College, Surampalem, India
e-mail: mudumbaisunaina@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_24

HAAR algorithm is used to detect the face in the frame and set the points, and this is implemented by using the OpenCV module in Python [2]. The main aim of proposed system is to provide an experience of virtual try-on with two different modes. Where one is to display ornaments on the image uploaded by the customer, and the other is to display the selected piece of jewellery on the once live preview [3]. The choice of selecting any of the modes is completely dependent on the customer's mindset. After selecting the mode, it will be considered as a frame, and it identifies the size and points of the face [4]. By using the face recognition library, it will fit the selected ornament on the customer's face [5]. This is achieved by the two main points in the application, they are overlaying faces with ornaments and detecting multiple faces.

1.1 Overlaying Faces with Ornaments

Overlaying face with ornaments is one of the main concerns in the application; this can be done by going through three important steps.

1. Find a face on the frame: To detect the face on the frame, OpenCV is used to identify the face in the provided frame irrespective of the given modes.
2. Detect specific landmarks on a face: Landmarks on the face are recognized by using an advanced library called face recognition with bounding boxes around the faces in the frame and with 68 facial landmarks points where the ornaments to be overlaid [6].
3. Overlay the image with an ornament: To go for the final step of overlaying the ornament first, the detected face is resized to get a little richer face region then a duplicate of the ornament is created with grey mesh for more accuracy in the detection. Finally, by taking the facial points, jewel is superimposed onto the customer's face.

1.2 Detecting Multiple Faces

In support for multi-face detection in the system, HAAR cascade is utilized where it is an object detection algorithm to identify faces in an image or a real-time video [7]. This algorithm uses edge or line detection features.

2 Related Work

Literature on multiple papers related to virtual overlaying of ornaments is done. The proposed system supports the user to try on the jewellery virtually with just a few steps. Firstly, the customer needs to face the camera, which adjusts the image

and overlay various articles of clothing and fabric to it and shows. This is done by using hardware components and expensive sensors. So without much expenses, a Web application is designed that displays the jewels virtually on the users face [8]. These are some of the existing applications with the idea of virtual try on of different articles.

- 3D form of fabric and purchaser's body using Asus Xtion sensor along with the virtual display [9].
- Using images of real garments allows for a realistic rendering quality with high performance [10].
- Ray-Ban uses a virtual mirror that helps the purchaser to virtually try sunglasses of the different frame before the purchase [11].
- DRAPE to simulate 2D clothing designs on 3D figures in different patterns and postures [12].
- 3D modelling of jewellery from the uncalibrated images [13].
- AR virtual fitting room which uses the Kinetic sensor as a 3D full-body capture device [14].

3 Proposed System

The proposed system provides interaction between user and augmented reality where user can try ornaments virtually. Everyone wants to try expensive, non-damaged jewellery. In this application, users can try jewellery virtually to reduce the physical contact risks such as ornament damage, robbery cases, and also increase the number of ornaments to try within the stipulated time. Users can try virtually live or by uploading photos of them. The ornament image will superimpose to the picture of user photo uploaded or live video. Based on the size and position, the jewellery will place on the image or live videousing OpenCV library. Using face recognition, face is detected in a frame in identify facial features such as eyes, chin, nose bridge, nose tip, eyebrows, and lips. By using, face recognition the jewellery object can be superimposed on the desired part. Python Flask Web Application interface is used to interact with the application.

The applications provide different feature and gender selection. Based on gender-specific distinct jewellery are shown for both men and women. Then, the users can select any jewels and try them on virtually. By clicking on the image which the user desired, it will redirect to another page where the result will display. Either picture uploading or the live user should select jewellery to try. The jewellery will be superimposed using the HAAR algorithm. HAAR algorithm is an effective way to detect objects. This algorithm, proposed by Paul Viola and Michael Jones in their paper rapid object detection using a boosted cascade of simple features. This algorithm has positive and negative images to train the classifier. Here, negative and positive images are

- Positive images: This contains the images which we want our classifier to identify.
- Negative Images: It doesn't have the object which we want to detect.

3.1 Advantages and Disadvantages of the Proposed Method

Advantages

1. It reduces the time, browses through the wide range of options.
2. Maximize selling by providing customers with a variety of options to choose from.
3. In this COVID situation, to avoid the spreading of infection in offline shopping this product is used so that users will not touch the jewellery.
4. Without touching jewellery, users will check whether a particular ornament will suit them due to this it minimizes the damage of jewels, i.e., prevents wear caused due to repeated trial.
5. The virtual ornament room gives a clear idea of the accessories which the customer wants to buy as a result providing better customer satisfaction.
6. Users can experience different models as per their request.
7. The application is user-friendly anyone can manage the application.
8. This application used AR technology so that the ornaments are realistic to use and select by the models and samples.
9. Virtual ornament room is roughly equivalent to the quality of the physical experience so there is no doubt in the designs of the accessories.
10. It can also take more than two persons or group photos for superimposing the ornaments.

Disadvantages

1. If there is no camera on the device, the user cannot experience the ornaments virtually.
2. It is not offline-based, there should be an Internet connection.
3. The user should upload a side picture see the real design of the nosepin.

Fig. 1 It shows the trial page where the user can select their gender

Fig. 2 It shows all the available jewellery for Tryon

According to the papers [2, 7, 15], the implementation of the virtual try-on is only limited to the women jewels. The proposed system stand apart from the other systems by providing gender selection feature and the try-on mode for men with the provided men jewels.

4 System Architecture and Working

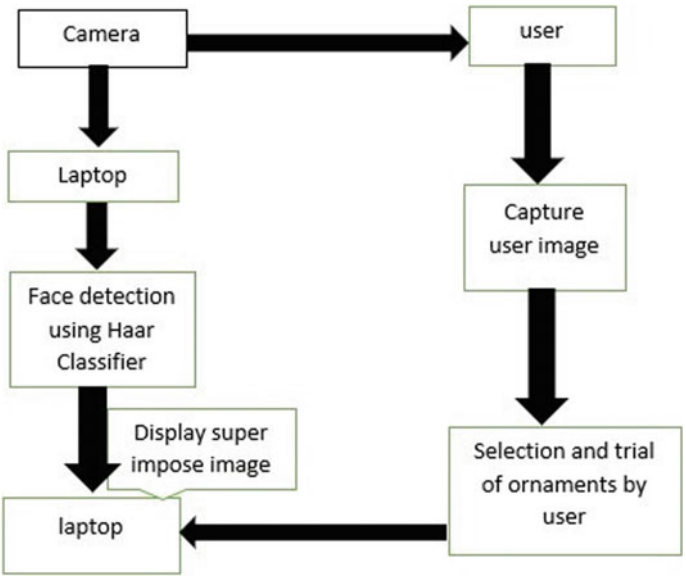


Fig. 3 System Architecture

Here, in Fig. 3, the camera means either upload the picture or live to stream. If you upload a file, it will detect the face using the Haar algorithm. Then, it will superimpose the jewellery image with the user's image. If you want a live stream, then the user should click on the live button. Then, it will capture the user's image, then the user should select the jewellery, and it will be displayed by superimposing the jewellery with the user's live video.

4.1 Custom Algorithm

Step 1: For the uploaded image, it identifies the landmarks using the face recognition library. It checks the length of face points and places x, y coordinates on the frame where the jewellery will be added
Step 2: For the selected ornament, it will calculate height and width
Step 3: All pixels greater than 230 will be converted to white, and others will be converted to black
Step 4: For the selected ornament, it will convert the background of the jewellery image to black
Step 5: Crop out the jewellery area from the original frame

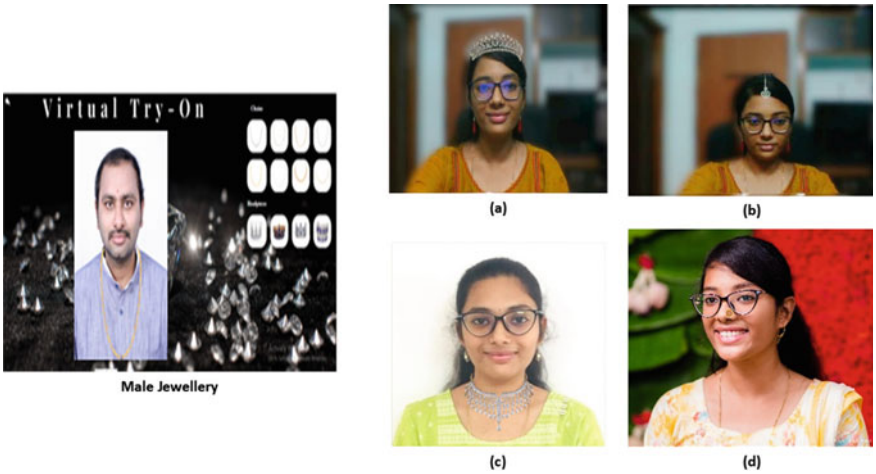


Fig. 4 Virtual Try of object crown, headpiece, necklace, and nose pin

Step 6: Bitwise AND operator will convert all black regions in any image to black in the resulting image and add both images so that the black region in any image will result in another image where as non-black regions being rendered over that area
Step 7: Replace original frame jewel area with newly created jewel area. Convert image to RGB format to showcase the output.

4.2 HAAR Algorithm

Step 1: Select the location from the facial Landmarks
Step 2: Crop a sub-image with the chosen pixel as the centre from the first image with similar dimensions as the convolution matrix
Step 3: Compute element-wise product for the values of a matrix and sub-image
Step 4: Add the result of the product
Step 5: Insert the resultant value into the new image at a similar location where you selected the pixel location.

5 Results

Figure 4 displays the feature provided for the men with the given jewellery and the result of the application in which the virtual object crown headpiece, necklace and nose pin are superimposed on live and the uploaded image of the user. This application is implemented by using OpenCV libraries with a HAAR algorithm for AR marking.

6 Conclusion and Future Scope

In conclusion, a virtual ornament room was executed successfully in Python OpenCV. This application helps purchasers to try jewels in less than a minute and sound safe while purchasing. It saves the time of going to the shops to try on the jewellery and gives a great level of satisfaction with a virtual try-on experience for the user.

The scope of the application is expanded by adding the other set of jewellery types like earrings, finger rings, bracelets, and different jewels models equally for men and women. By designing the application as a plug-in, it can be made available to any online e-commerce jewellery site for providing a virtual try-on experience for the customers.

References

1. Soo S (2014) Object detection using haar-cascade classifier. Inst Comput Sci Univ Tartu 2(3):1–12
2. Salunke G, More H, Shete R, Kawade P, Virtual jewellery shopping using augmented reality. Int J Eng Res
3. Cuimei L, Zhiliang Q, Nan J, Jianhua W (2017) Human face detection algorithm via haar cascade classifier combined with three additional classifiers. In: 2017 13th IEEE international conference on electronic measurement & instruments (ICEMI), IEEE, pp 483–487
4. Sanivarapu PV (2021) Multi-face recognition using cnn for attendance system. In: Machine learning for predictive analysis. Springer, pp 313–320
5. Kasinski A, Schmidt A (2010) The architecture and performance of the face and eyes detection system based on the haar cascade classifiers. Pattern Anal Appl 13(2):197–211
6. Kim HW, Kim HJ, Rho S, Hwang E (2020) Augmented emtcnn: a fast and accurate facial landmark detection network. Appl Sci 10(7):2253
7. Thakur A, Virkar S, Gaikwad J (2020) Online virtual trial room implementation using opencv python. Int Res J Eng Technol
8. Saibaba G, Sanivarapu PV (2018) Developing an userfriendly online shopping web-site. Indones J Electr Eng Comput Sci 12(3):1126–1131
9. Paolanti M, Pietrini R, Mancini A, Frontoni E, Zingaretti P (2020) Deep understanding of shopper behaviours and interactions using rgb-d vision. Mach Vis Appl 31(7):1–21
10. Hauswiesner S, Straka M, Reitmayr G (2013) Virtual try-on through image-based rendering. IEEE Trans Vis Comput Graph 19(9):1552–1565
11. Baek TH, Yoo CY, Yoon S (2018) Augment yourself through virtual mirror: the impact of self-viewing and narcissism on consumer responses. Int J Adv 37(3):421–439
12. Guan P, Reiss L, Hirshberg DA, Weiss A, Black MJ (2012) Drape: dressing any person. ACM Trans Graph (TOG) 31(4):1–10
13. Radhika G, Kumar TS, Reyner PPD, Leela G, Mangayarkarasi N, Abirami A, Vinayaka K (2016) 3d modelling of a jewellery and its virtual try-on. In: Proceedings of the international conference on soft computing systems. Springer, pp 157–165
14. Pachoulakis I, Kapetanakis K (2012) Augmented reality platforms for virtual fitting rooms. Int J Multim Appl 4(4):35
15. Zhang T, Wang WYC, Cao L, Wang Y (2019) The role of virtual try-on technology in online purchase decision from consumers' aspect. Internet Research

Performance Analysis of an ADM Ring Network for Optical Communication System

Amit Grover, Shyam Akashe, Anu Sheetal, Reecha Sharma, and Meet Kumari

Abstract In this paper, the performance of the ADM optical ring network is demonstrated. The four nodes ring network at 40 Gbps is demonstrated. It is analyzed that the signal quality rises as the number of nodes increased. It is found that the signal can be transferred up to a communication distance of 300 km with acceptable BER and Q factor. The degradation of a signal occurs after 300 km distance. The received power at various input powers is investigated, and results are also calculated by eye diagrams from which crosstalk of a signal can be observed.

Keywords ADM · Quality factor · BER · MZM · SFM

1 Introduction

In optical networks, there is a probability to upgrading the current networks to WDM networks due to large bandwidth and flexibility. In practical systems, the number of signals can mix with each other and cause crosstalk in the network. They increase signal degradation and a bit error probability [1, 2]. Add–drop multiplexers are the

A. Grover
Department of Electronics and Communication Engineering, Shaheed Bhagat Singh State University, Ferozepur, Punjab, India

S. Akashe
Department of Electronics and Communication Engineering, ITM University, Gwalior, M.P., India

A. Sheetal
Department of Engineering and Technology, Guru Nanak Dev University, Regional Campus, Gurdaspur, Punjab, India

R. Sharma
Department of Electronics and Communication Engineering, Punjabi University, Patiala, Punjab, India

M. Kumari (✉)
Department of Electronics and Communication Engineering, Chandigarh University, Mohali, Punjab, India
e-mail: meet.e8072@cumail.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_25

key module in optical networks to monitoring the movement of data. ADM increases network capacity and flexibility. They drop and add any wavelength at an intermediary node deprived of disturbing another channel, and dropped wavelengths can be reused [3]. It is hard to improve the component factors for a least crosstalk due to altered sources of crosstalk. Crosstalk in ADM is affected by internal reflections, switch leakage, and multiplexer/demultiplexer outflow. Interchannel crosstalk is arising from an ADM or optical filter that chooses one channel and passes others over the network. The switching between different wavelengths produces crosstalk from the defective isolation among the switch ports. Other crosstalk can be produced by amplifier, router, and cross-phase modulation. A signal can accrue crosstalk from various channels and elements among the network. Several techniques can be used to reduce crosstalk such as filter cascading or wavelength dilation [4–6]. Ring topology is selected due to less number of links. Ring networks share the load and have better flexibility. Each node connects to other node and creates a single unbroken path for signals over each node [7]. The data travel in the form of packets from node to node. Rings can be bidirectional or unidirectional, and all the data are traveling either clockwise or anticlockwise direction. A unidirectional ring network topology delivers one path among any two nodes and disturbed by the breakdown of a particular connection. A cable break and a node failure may separate each node involved to the ring [8]. Kumar et al. [9] analyzed the interchannel crosstalk at add–drop multiplexers of a ring network at sweeping filter bandwidth. A ring network has four nodes at 10 Gbps over two channels at 1552.2 nm. The distance between different nodes is 12.5 and 25 km, and an amplifier is placed before one of the nodes to recompense losses. Bhatia et al. [10] demonstrated the crosstalk using eye diagrams at ADM ring network at sweeping filter bandwidth and showed that there was no signal deprivation in the ring network. The signal retains on refining as it passes over the successive nodes. It was analyzed that by increasing the signal power, more number of users can support. The investigation presented in [9, 10] limited to transmission distance and data rate. The earlier work is improved by longer distance, and the system performance is demonstrated at 40 Gbps data rate. The organization of paper is into the following sections: First section describes the introduction; the system description for the communication system is described in Sects. 2 and 3 which labels the result and discussion. Final, conclusion is presented in Sect 4.

2 System Description

Figure 1 shows the architecture of an optical ADM. ADM comprises an optical demultiplexer, an optical switches, and an optical multiplexer. The optical demultiplexer splits the signal in an input fiber onto ports. The optical switches are used for reconfiguration and pass the signal to the multiplexer or to drop ports.

The data from the input side and add ports both are demultiplexed. These signals are exchanged independently using a single switch. The incoming data are directed to the output mux or the drop mux depending on the switch state. The data are then

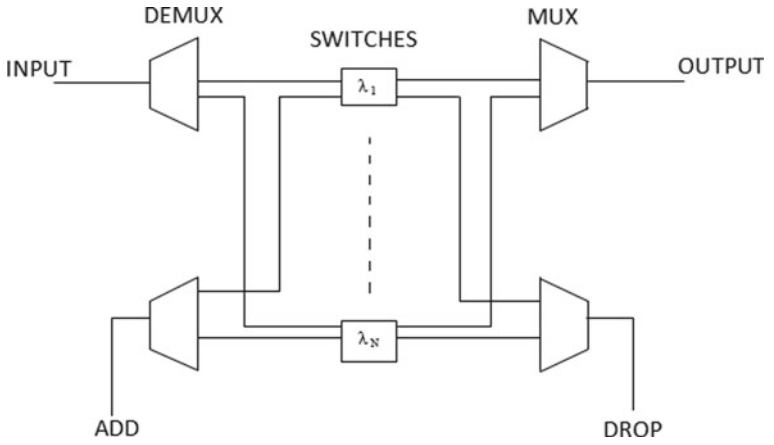


Fig. 1 Architecture of a reconfigurable ADM

remultiplexed. Incoherent and coherent crosstalk will rise due to drip in the switches, muxs and demuxs. If a signal leakage arrives at the output port, then the incoherent crosstalk will be detected on the channel because the two channels are produced by different lasers, and the phase difference varies randomly. Thus, the interference produces noise. The interference is exploited if both the fields ensure equal state of polarization. The noise is called phase-induced intensity noise. Coherent crosstalk arises as an effect of interference amid a signal and a delayed signal.

The crosstalk effect on a WDM ring is shown in Fig. 2. The network contains N number of nodes, and all are transceivers. The transceiver contains a PRBS generator, a CW laser, a Mach–Zehnder modulator, ADM, an attenuator, a photodetector, and a LPF. All transceivers are connected through a single mode fiber. Erbium-doped fiber amplifier with noise figure 4 dB and gain 10 dB has been placed before node 2. The bandwidth of transceivers is 10 GHz. The distance between the nodes is 30 km. At

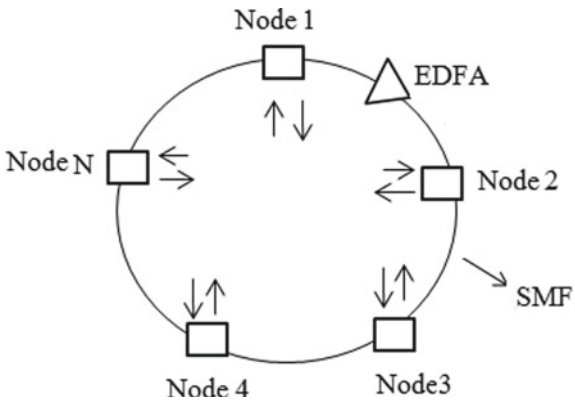


Fig. 2 Architecture of a ring network with four nodes

every node, one frequency is added and other dropped. The frequency f_1 193.1 THz is added at node 1, and the frequency fn is dropped. All the wavelengths are recycled in a round way in a ring, i.e., the freq f_1 is added, and the freq f_4 is dropped at node 1. The freq f_2 is added, and the freq f_1 is dropped at node 2 and so on. If the number of nodes is enhanced in the network, the ring goes so on.

3 Results Analysis

The performance of the system is defined in terms of quality factor, bit error rate, and eye diagrams. It is seen from Fig. 3 when the ring organization retains on repeating, the quality factor is refining.

It is depicted that from the graph as the fiber length can increase up to 300 km, after that Q factor keeps on falling. As the length of fiber is increased, enhancement can be analyzed up to a certain range. Also, as the fiber length increases from 50 to 300 km, the value of Q factor is reduced to some extent. The result shows improvement over [9, 10] with transmission distance of 300 km.

Figure 4 shows that least BER is increased to some extent for different nodes as the distance is increased from 50 to 300 km. It is analyzed that with the enhancement in fiber length, performance of the system declines and crosstalk rises.

The more users can support by enhancing the input power. As the signal input power is increasing, enhancement in output optical power is detected. The output power also rises as the input power rises for entirely users. The power consequence

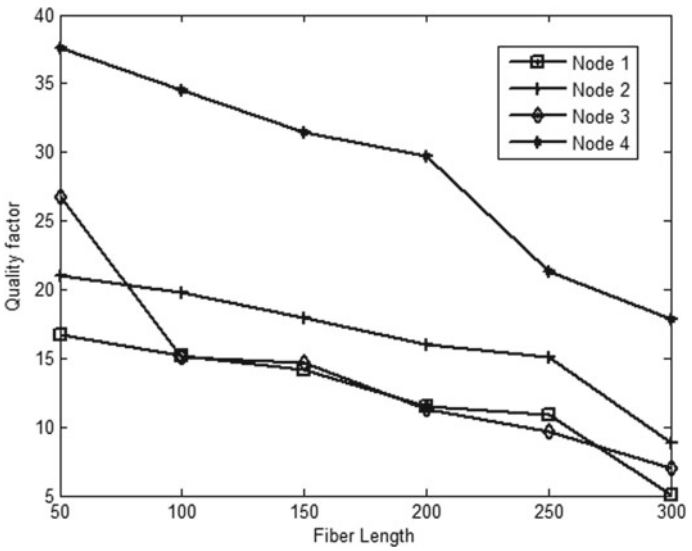


Fig. 3 Quality factor vs. fiber length at different nodes

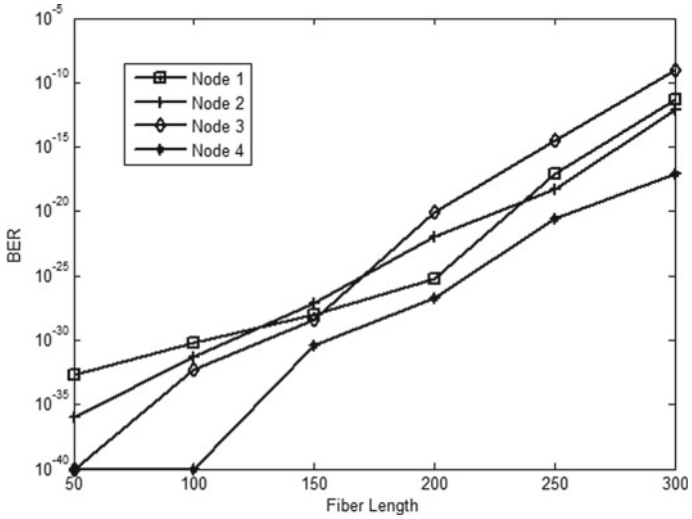


Fig. 4 BER versus fiber length at different nodes

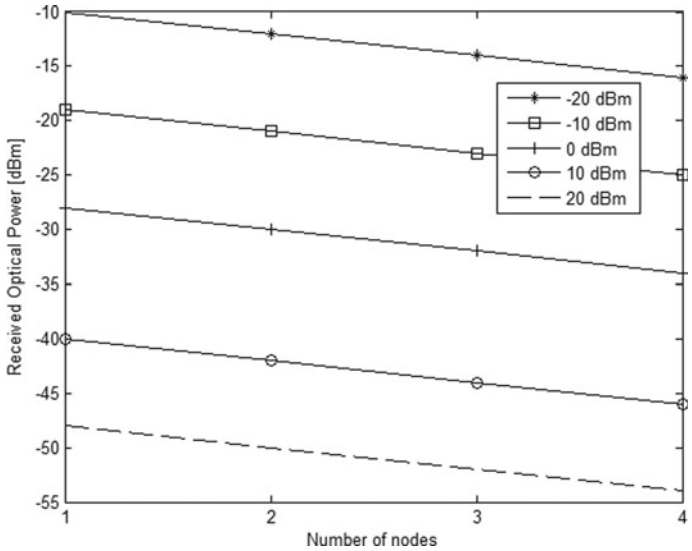


Fig. 5 Received optical power vs. number of nodes at different input optical power

goes on accumulative with the enhancement in the input power. When the power retains on accumulative up to 20 dBm, the output power also enhances. Therefore, large number of users can support by enhancing the power up to a definite range. This is due to the rise in input power; the received power rises. Figure 6 shows the eye diagrams from which several signal factors can be observed.

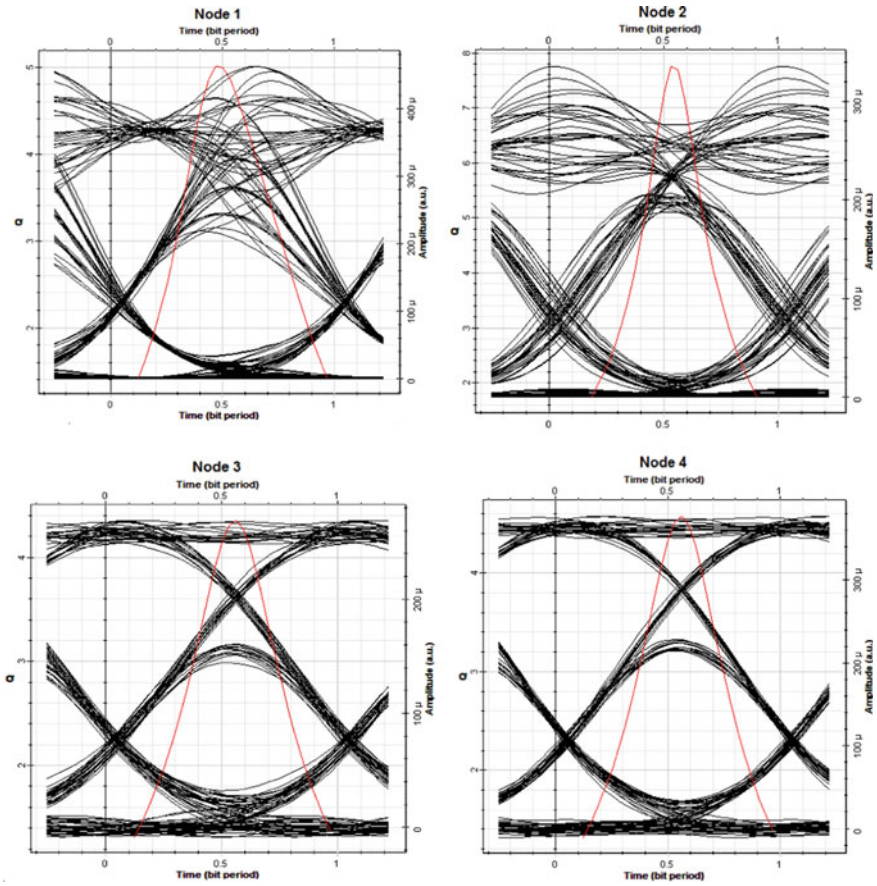


Fig. 6 Eye diagrams for different nodes

It is found that there is proliferation in quality of the signal taken at the fourth node. The consequence of noise is more on the initial users. The eye opening is much wider at node 4.

4 Conclusion

The performance of ADM optical ring network with 40 Gbps bit rate is evaluated. It is observed that the acceptable Q factor and bit error rate are achieved up to 300 km transmission distance. The demonstration of the system declines to some extent and crosstalk increases. When the input power rises up to 20 dBm, the output optical power rises. Therefore, more users can support by enhancing the input power up to a

definite range. The dispersion effects and crosstalk became more conspicuous with the increase in the optical fiber length.

References

1. Gyselings T, Morthier G (1999) Crosstalk analysis of multiwavelength optical cross connects. J Lightwave Technol 17(8):1273–1283
2. Dewra S, Kaler RS, Placement of dynamic reconfigurable optical add/drop multiplexer using dynamic channel equalizer in an optical communication system. Optoelectron Adv Mater Rapid Commun 9(1–2):53–56, Jan–Feb 2015
3. Mukherjee PP, Sarkar S, Das NR (2014) Analysis of in-band crosstalk-crosstalk beat noise and obtaining Bit Error Rate in a WDM receiver. In: IEEE international conference on automation, control, energy and systems (ACES), 1–2 Feb 2014
4. Kim CH (2012) Dependence of in-band incoherent crosstalk-induced penalty on seed source in RSOA-based WDM-PONs. IEEE Photonics Technol Lett 24(7):581–583
5. Dewra S, Kaler RS (2013) Crosstalk analysis in an optical network based on optical cross connects with different MZI techniques. Optik 124:55–59
6. Goldstein EL, Eskildsen L, Elrefaie A (1994) Performance implications of component crosstalk in transparent light-wave networks. IEEE Photon Technol Lett 6:657–659
7. Sajeeb AA, Farin FT, Rahim A, SP Majumder (2016) BER performance limitation of an optical link with multicore fiber (MCF) due to crosstalk caused by mutual coupling among the cores. In: 3rd international conference on electrical engineering and information communication technology, IEEE
8. Kang C-S, Park B-S, Shin J-D, Jeong J-M (1995) A broadband ring network: multichannel optical slotted ring. Comput Netw 27(9):1387–1398
9. Kumar V, Sharma AK, Agarwal RA (2010) Investigation on interchannel crosstalk at ADM for an optical ring network. Optik 121:45–49
10. Bhatia B, Shokeen V, Singh R (2012) Evaluation of optical ring network to measure inter-channel crosstalk using “eye diagrams”. Int J Comput Appl 40(15):35–40, Feb 2012

Biometric-Based Authentication in Online Banking

Shivaraj Hublikar, Vishal B. Pattanashetty, Venkatesh Mane, Preeti S. Pillai, Manjunath Lakkannavar, and N. S. V. Shet

Abstract The use of technology has become the integral part of human life. Most importantly the introduction of Internet has made the lives of people easy and due to its cost effectiveness; usage of Internet by people has been increased. The Internet has made the people to move toward the online mode of transaction. Banks recommend their customers to use Internet banking facility and assure it as a safest mode of transaction but it is associated with huge risk. The continuous rise in online banking brings several security issues and increased cost of implementing higher security systems for banks and customers. Present online banking technology works on 2-Factor authentication mode, i.e., it works on both transaction level and authentication level. The main problem associated with 2-Factor authentication is that there may be risk associated, like cyberattack (SIM swapping fraud), etc., so we have proposed a biometric authentication technique which makes transaction more secure.

Keywords Authentication · One time password OTP · Fingerprint · E-banking

S. Hublikar (✉) · V. B. Pattanashetty · V. Mane · P. S. Pillai
K.L.E. Technological University, Hubli, India
e-mail: shivaraj@kletech.ac.in

V. B. Pattanashetty
e-mail: vishalbps@kletech.ac.in

V. Mane
e-mail: mane@kletech.ac.in

P. S. Pillai
e-mail: preeti_pillai@kletech.ac.in

M. Lakkannavar
MSRIT, Bengaluru, India
e-mail: manjunathl@msrit.ed

N. S. V. Shet
NIT-Karnataka, Mangaluru, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_26

1 Introduction

The advancements in the computing standards and Internet banking has changed significantly the way one performs banking transaction. Since the banks have adopted online mode, it has become easy and faster for the people to make transaction online rather than going to bank and waiting in queue. This made the people to use online mode widely, but this method is not completely secure. As the number of users and cyber-crime increased it was essential for banks to make their application more secure for which they used 2-Factor authentication, i.e., sending OTP to the user's registered mobile numbers. But there are some loopholes in 2-Factor authentication. In modern days bank requests customer to link their mobile number, so that they can use online banking services seamlessly. Here the main problem is fraudster can take duplicate SIM of customer from network operator by providing necessary government proofs and install that SIM on their mobile and they can have access to customer's bank account and they may use it. This is known as SIM Swapping Fraud. So to address above problem this paper gives appropriate solution. In this paper biometric authentication is used to login, i.e., the user receives the OTP which will be in the encrypted form here encryption and decryption takes place with the help of AES Algorithm. To make transaction user needs to verify his finger print so that the encrypted OTP gets decrypted and transaction gets successful.

1.1 Literature Survey

DTI information security breaches survey 2006 revealed that still the companies businesses depend on user IDs and password to verify the identity of the users who are trying to access their systems. Single factor authentication is being used by many companies as it uses single static password which has the benefit for user to remember it easily. But these static passwords can be cracked easily and will be vulnerable to user. Changing the password regularly is more secure than static passwords. A disadvantage of changing passwords regularly is that they can be easily forgotten. OTP is password which is valid for some time and after that it dies. As the static passwords are vulnerable to attacks, along with static passwords, OTP is used to login or making transaction [1]. Here user gets encrypted OTP, then user types the application password and send it to the Database server there the OTP matching takes place and proceed to next step which gives authentication for the system.

In present process of online banking, there are two types of authentication, i.e., user authentication and another one is data origin authentication. Presently, there are various attacks like mitb4, esa, Zeus, etc., so [2] they are making use of quick response codes (QR code). Some of the most important features of this are resolution, Data type, and Error correction. By using QR codes in e-banking we can reduce the attacks. Here we will be having a decoder which we decode the QR. By implementing

this we can reduce security measures and it will be an efficient in modern banking systems.

As today online mode is growing and has become global way to socialize, transact, and communicate, so it has become essential to make things more secure and there are many ways to make online mode more authentic like using biometric iris reading, gestures or voice recognition, pattern-based login, etc. But still login/password pair is widely accepted method to authenticate. It is true that there is flaw in all these methods. This paper proposes different idea on authenticating online mode like using e-biometric authentication signature which is developed using SHA1, RSA algorithm, using Palm Vein Technology, using Personal Identification Number, using authenticating smart card. But the proposed technique is to merge two authentications, i.e., conventional PIN-based authentication and second is the palm vein technology [3].

Online authentication is to identify and verify through Internet. As now HTTPS is being widely used for online banking which contains Hyper Text Transfer Protocol and secure socket layer. This authentication process starts with; first bank provides proof by sending digital certificates issued by certificate authority. This certificate contains information like issuer's name validation time or period of certificate, serial and version number etc., then client or customer confirms and provides his credential to bank and then bank verify the identity of client and starts communication under symmetric encryption. It also provides the client with transaction confirm functionality to prevent fake transfer of capital [4]. Most of the banks are recommending their customers to use Internet banking and assure it as safest method. Customer uses the application of bank for online transaction. During the transaction the customer receives the OTP which is received from the Database server then customer enter the received OTP for successful transaction, but the problem lies here is that one's mobile can be misused by anyone. So to tackle this problem the verification finger print during transaction is done [5].

In [6] architecture (security) of Short Message Service (SMS) systems and threats toward it is analyzed. The basic foundation on which SMS architecture is on is cellular network and mobile hardware. In earlier days SMS played a keen role, it was brought into practice to reduce phishing attacks against authentication and authorization of e-banking services. A virtual dedicated channel (OTP) has been designed to facilitate SMS and it requires minimal change in mobile hardware. The designed channel gives protection against Trojans and malware attacks.

In [7] it deals with QR code-based phishing attacks. There are different in Internet banking like Trojan, phishing, masquerading, etc. which captures important data of user and use it in making frauds. QR code was built in Japan to use in the automotive industry. We can feed any type of data in QR code like plain text, SMS link, URL, etc. In this steady research they have been able to detect and design a novel game design framework with qualitative and quantitative analysis. At last they are able to enhance user's behavior by encouraging them from phishing attacks.

There are lots of way to make the system secure like traditional way of user ID and password, and OTP, photo matching, biometric ways, etc. But if we see biometric way is more secure and good way to protect data. Example finger print

authentication, as we can see no two human being has same finger print. Finger print has many detail like crossover, core, delta, pore, ridge ending and many more things and these differs person to person. Earlier this used to be done using ink manually, but nowadays optical sensor, ultrasonic and solid state sensors are used to check if owners fingerprint match with scanned ones or not [8].

In [9] case study on Quick Response (QR) code has been studied deeply by selecting Library as the field. QR code is 2-D matrix bar codes. It was developed by Toyota subsidiary Denso wave. QR codes are basically in image format like JPEG, BMP, and PNG, etc. Some of the functions of QR code are encode phone numbers, send a v-card, bookmark a site link, etc.… A QR code can store a maximum of 100 fold of information. It has self-error correction capacity. These QR code can be used in libraries in the form of bar codes for easy searching of books.

In this paper everything about fingerprint is mentioned. This paper has proposed a new and improvised idea for fingerprint authentication or identification. As nowadays biometric way of authentication is growing it becomes essential to make it more secure. The solution of identification is done by measuring Euclidian distance from center point to the nearest bifurcation and ridge ending minutiae. First the finger print is scanned and stored in a database and then during verification time its Euclidian distance is measured [10].

1.2 Organization of the Paper

In this section, we discuss about the organization of report. It includes system design, implementation details, results and discussion. System design includes the methodology adopted the algorithms used and flowchart. Implementation details include the proposed process flow. Results and discussion show the results obtained during our work.

2 Methodology

Traditional Process Flow

In traditional flow when user request for OTP in mobile application, the OTP gets generated and received in user's mobile phone. The received OTP will be generated from Database server where all users' details will be stored. After receiving OTP user enters the OTP and sends back the OTP for verification to Database server. If the OTP is verified then user's transaction will be successful.

AES Algorithm

AES was established in 2001 by the US NSIT. As key used in DES algorithm has small size and low computing capability AES algorithm was developed. It has a

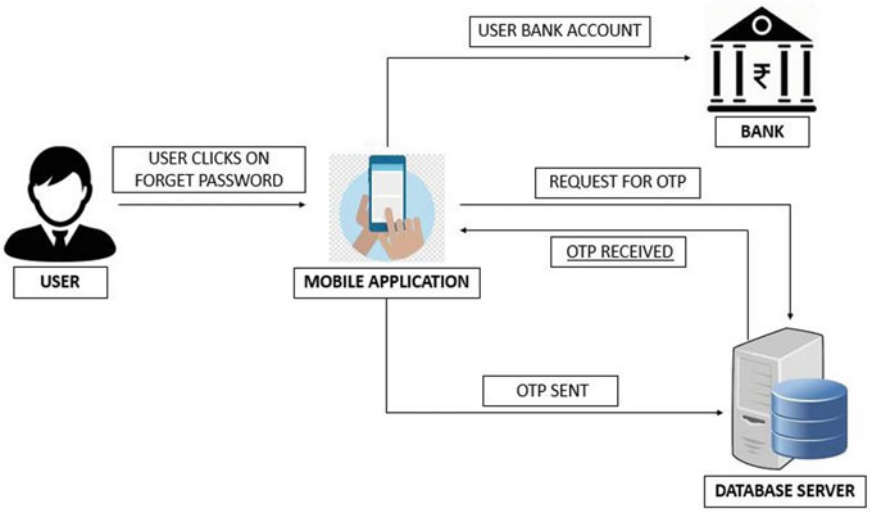


Fig. 1 Normal process flow

symmetric key block chain, means same key is used for encryption and decryption. AES algorithm is six times faster compared to DES algorithm. The number of bits used in key depends on number of rounds (i.e., iteration). Ten rounds take 128 bit key that is why it is called AES 128 version. Similarly for 12 rounds it takes 192 bit key and for 14 rounds 256 bit key. As seen in Fig. 1 the block Key Expansion Algorithm which expands cipher key of 128 bit. The number of keys generated in key expansion is one more than number of rounds. In the initial stage only XOR operation is done between input data and key of four words. Each round except last round uses 4 transformations, last round has only 3 transformation, i.e., substitution bytes, shift rows, add Key. In decryption cipher keys are applied in reverse order.

Input plain text is in the form of 4×4 matrixes which is equivalent to 16 bytes and this 16 bytes data passes through four transformations as shown in the Fig. 2. This process continuous for nine rounds, but for the tenth round mix-column transformation does not happen. At the end of tenth round output will be 128 bit cipher text (Fig. 3).

Database Server

Google Fire base platform is used as a Database server to store user data and authenticate the user in the time of login. User data such as Email, Account number, and Bank balance of the user is stored in fire base and can be fetched once the user verifies his fingerprint. To store the data of any user in fire base his account number is used as a key and if a user wants to fetch/modify the data then firstly account number of the user is generated internally based on his login details, i.e., Email and is searched in fire base where all his details are stored as a child to this key (Account number) and s

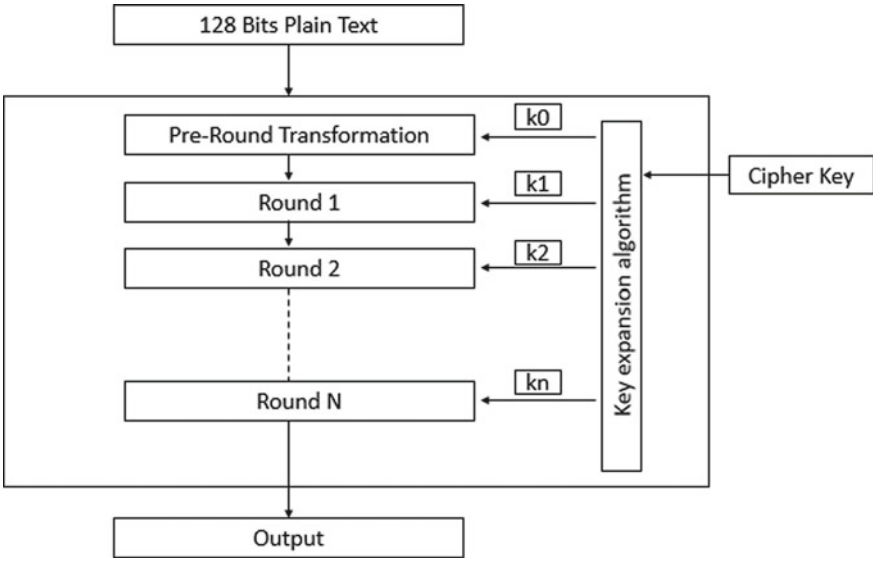
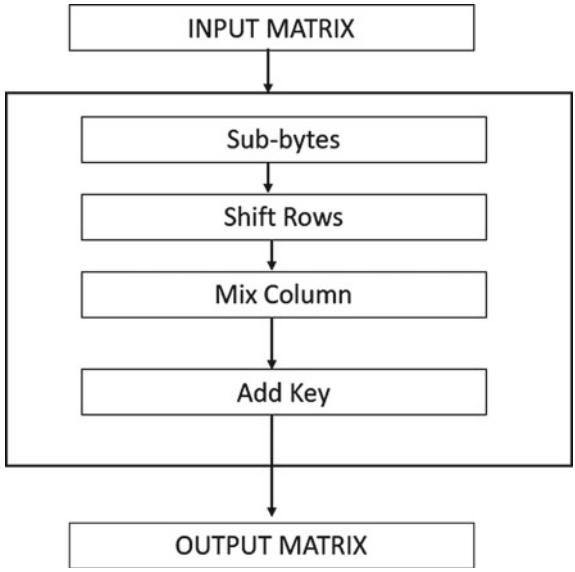


Fig. 2 General design of AES encryption

Fig. 3 Structure of each round at the encryption side



fetched and displayed to the user. Also there is a "forgot password" option available for the user to change the password if he forgets it.

Flow Chart

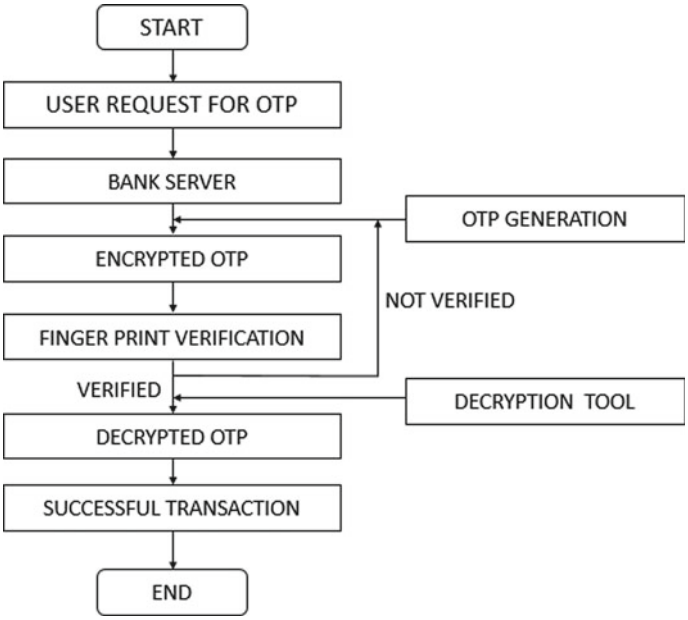


Fig. 4 Flow chart

3 Implementation

In proposed process flow show in Fig. 4 when user request for OTP in mobile application, the OTP gets generated and received in user's mobile phone. The received OTP will be in encrypted format which is received from Database server. Now user needs to verify his finger print for decryption of OTP, if the verification of OTP is successful then OTP decryption will be successful and thus user's transaction will be successful. This technique can also be used to check bank balance of user (Fig. 5).

When user installs "Securing Bank OTP" application in mobile, first the user needs to create account in Database server by providing necessary information like email ID, account number, mobile number and user needs to set password. If user has already created account and if the user is trying to register using same information will not be possible. Every user will be having unique account number. Once the user has registered his/her account in database server, now user needs login into application through email ID and password that has been entered during registration as shown in Fig. 6. If user forgot password then user can reset password by tapping on forgot password option. There user needs to provide email ID and user gets reset password link to his registered email ID forgot password page (Fig. 7).

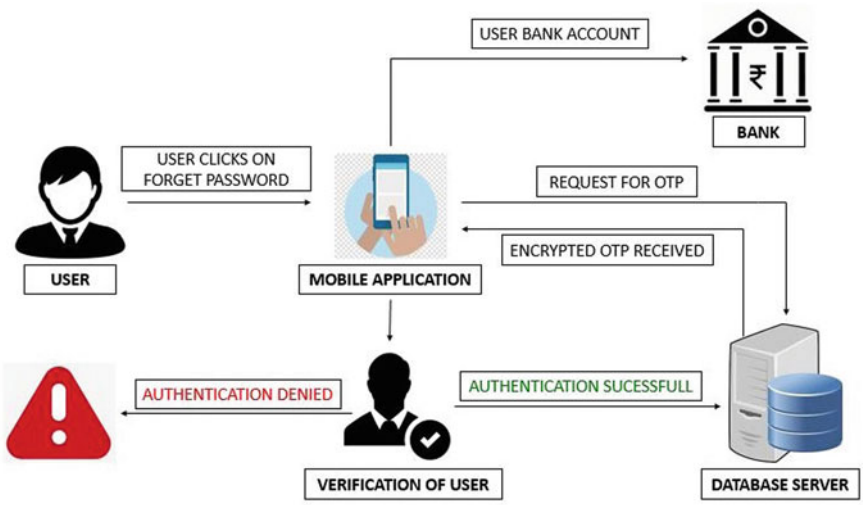


Fig. 5 Proposed process flow

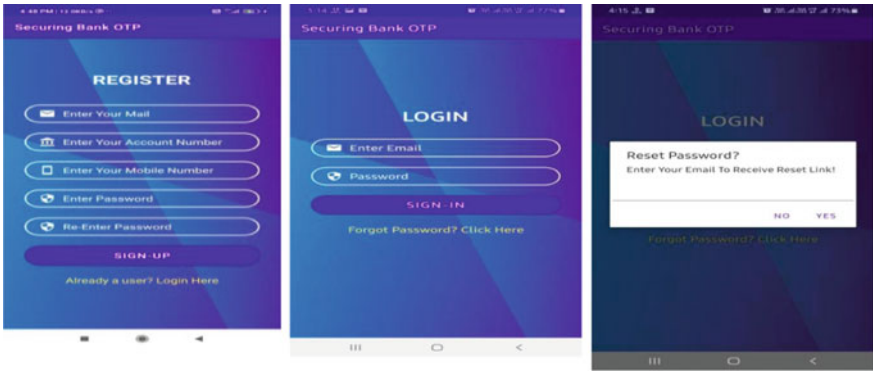


Fig. 6 Registration page for new user, Login page for registered user, Reset password page

4 Results and Conclusion

4.1 Android Application

If the user forgot the password, can be reset by tapping on forget password option in login page and user will get mail to reset the password. Once the user successfully logged into application, the user can have access to four services namely, to check account balance, to make transaction, Scan QR code and My QR code as shown in Fig. 8. The My QR code image can be shared through various social media handles. If user wants make transaction to his own account with his/her account it is not be

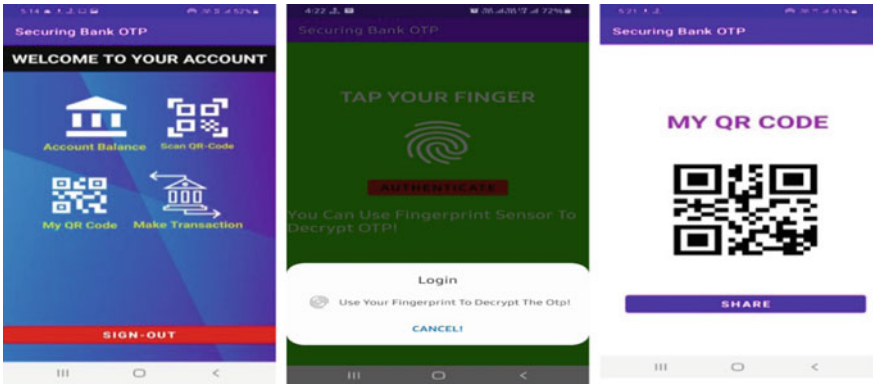


Fig. 7 User account interface, authentication page, My QR code page

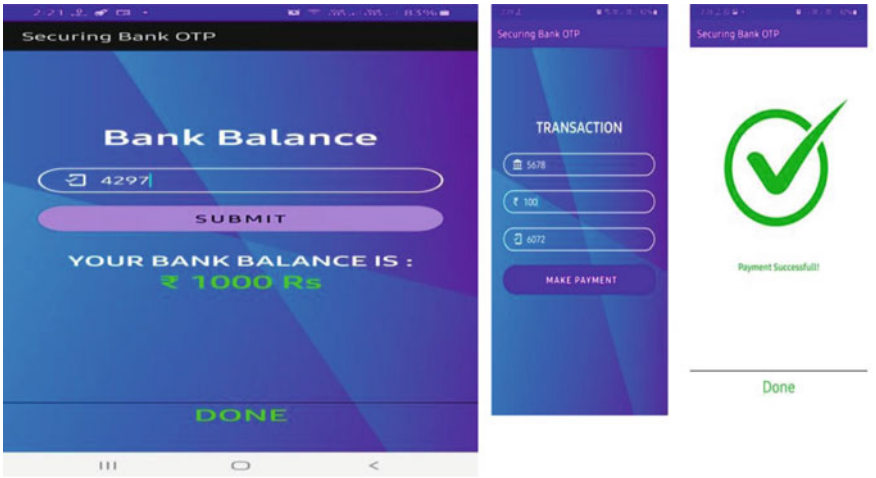


Fig. 8 User account balance, transaction page

possible. If user wants to check bank balance or to do any transaction, user needs to decrypt the OTP by verifying registered finger print. If user wants to share his account number with another user, then he can directly share his QR code image with them. At the receiver side he needs to decode the QR using scanner to get account number as shown in Fig. 8. The user will be able to check the balance by tapping on account balance option and then gets encrypted OTP to his/her phone and then user decrypt the OTP by verifying the finger and then enters the decrypted OTP to view the account balance as shown in Fig. 8. The user can have the option to transfer the money form one account to another account by tapping on Make Transaction option and then enter the account number of other user and then enter the money and also enter the OTP and proceed to make payment as shown in Fig. 8.

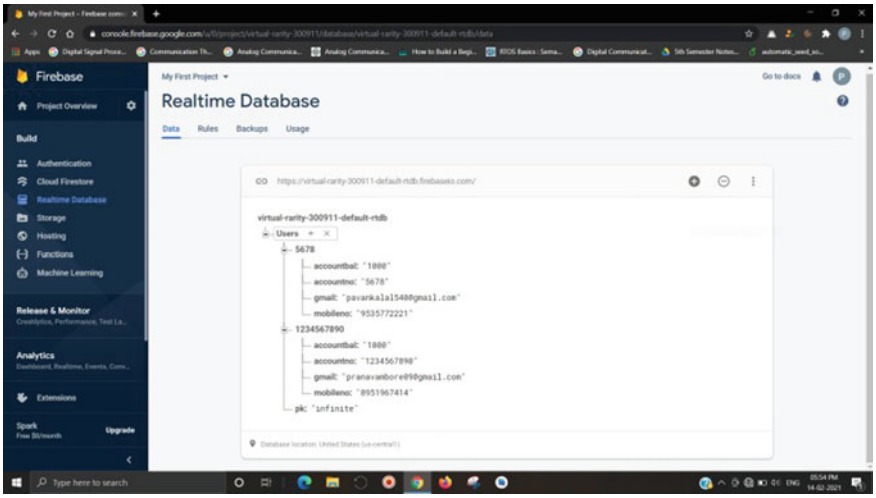


Fig. 9 Database of registered users in fire base

Google Fire Base

We are using Google Fire base to store all the details of the registered users. This acts as Database server of the bank. In this all the registered user details like Account Balance, Account Number, Gmail, and Mobile number gets stored (Fig. 9).

References

1. Kalaikavitha MCAE, J. Gnanaselvi, Secure login using encrypted one time password (Otp) and mobile based login methodology. Int J Eng Sci
2. Kumar D, Agrawal A, Goyal P (2015) CSE, graphic era university, "Efficiently improving the security of OTP". In: 2015 international conference on advances in computer engineering and applications (ICACEA) IMS engineering college, Ghaziabad, India
3. Mulliner C, Borgaonkar R, Stewin P, Seifert J-P, SMS-based one-time passwords: attacks and defense. Northeastern University crm@ccs.neu.edu 2. Technische Universit ¨at Berlin
4. Dayaratne TT, WSO2 Sri Lanka thusithathilina@gmail.com. "A framework to prevent QR code based phishing attacks". Research Gate
5. Balti A, Sayadi M, Fnaiech F, SICISI unit, ESSTT, University of Tunis, 5 Av. Taha Hussein, 1008, Tunis, Tunisia." Improved features for fingerprint identification". ieee.org
6. Pattanashetty VB, Iyer NC, Ibal P, Bagewadi A (2015) Smart driving assistance. In: 2015 IEEE international transportation electrification conference (ITEC), Chennai, pp 1-4, https://doi.org/10.1109/ITECIndia. 2015.7386870
7. Pattanashetty VB, Iyer NC, Viswanath HL (2018) Intellectual conveyance structure for travellers. In: Saeed K, Chaki N, Pati B, Bakshi S, Mohapatra D (eds) Progress in advanced computing and intelligent engineering. Advances in intelligent systems and computing, vol. 564. Springer, Singapore
8. Pattanashetty VB, K Shamshuddin, Iyer NC (2020) A top down approach from job to course. Procedia Comput Sci 172:204–206. ISSN 1877-0509

9. Pattanashetty VB, Iyer NC, Viswanath HL, Kore S (2019) Inclusive device to connect people with the surrounding world. In: Fong S, Akashe S, Mahalle P (eds) Information and communication technology for competitive strategies. Lecture notes in networks and systems, vol 40. Springer, Singapore
10. Pattanashetty VB, Iyer NC, Dinkar A, Gudi S (2017) Self-coordinating bus route system to avoid bus bunching. In: Satapathy S, Bhateja V, Joshi A (eds) Proceedings of the international conference on data engineering and communication technology. Advances in intelligent systems and computing, vol 469. Springer, Singapore

Various Diabetes Detection Techniques a Survey

Shahee Parveen, Pooja Patre, and Jasmine Minj

Abstract Diabetes Mellitus is one of the most serious illnesses, and it affects a large number of people. Diabetes Mellitus may be caused by age, obesity, lack of exercise, genetic diabetes, lifestyle, poor diet, high blood pressure, and other factors. Diabetics are at a better chance of developing such as heart failure, kidney disease, stroke, eye problems, nerve damage, and so on. The standard hospital procedure is to collect the necessary information for diabetes diagnosis is obtained through a variety of tests, and appropriate medication is prescribed based on the results. Type 1 and 2 diabetes are the maximum not unusual place sorts of the condition, however there also are different kinds, inclusive of gestational diabetes which happens during pregnancy, as well as other forms. The emphasis of this paper is on various prediction techniques that have had a major impact in the field.

Keywords Data mining · Genetic algorithms · Artificial neural network · Extreme learning machine (ELM) · Support vector machine (SVM) · Logistic regression

1 Introduction

Diabetes is also a chronic disease that spreads into the blood and a metabolic disease that has become increasingly common and widespread in India. Insufficient secretion or impaired insulin action are the causes. Diabetes mellitus is a category of disorders characterized by hyperglycaemia, or an abnormally elevated blood glucose level, as well as altered metabolism of carbohydrates, lipids and proteins, described by hyperglycaemia, deficient insulin, and insulin obstruction or both, prompting modified digestion of starches, protein, lipids and an expanded danger of vascular intricacies, imperfection in receptive oxygen species rummaging chemicals and high oxidative pressure-actuated harm to pancreatic beta cells. Diabetes is a constant illness with the possibility to cause an overall medical care emergency. As per International Diabetes Federation, 382 million individuals are living with diabetes across the entire world.

S. Parveen (✉) · P. Patre · J. Minj
Department of Computer Science and Engineering, VEC, Lakhanpur, CG, India
e-mail: Shaheeparveen14026@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_27

By 2035, this will be multiplied to 592 million. Different kinds of revealed diabetes mellitus can be characterized into two classes named as Type1 diabetes and Type2 diabetes.

- Type 1 Diabetes mellitus happens as an immediate consequence of the pancreas not assembling the glucose-controlling chemical. This specific type of diabetes was recently called “adolescent diabetes”.
- Type 2 Diabetes mellitus starts with protection from the pancreatic created chemical, and the cells of the body neglecting to react to the chemical. Movement of the illness may make an absence of the chemical likewise create. This type of diabetes was recently known as “adult-onset diabetes”.

1.1 Data Mining

For the most part, information mining (some of the time called information or information revelation) is the way toward examining information from alternate points of view and summing up it into valuable data—data that can be utilized to expand income, reduces expenses, or both. Information mining programming is one of a few insightful apparatuses for investigating information. Information mining is fundamentally utilized today by organizations with a solid purchaser center—retail, monetary, correspondence, and showcasing associations. It empowers organizations to decide the effect on deals, consumer loyalty, and corporate benefits. Information mining comprises of five significant components:

- Concentrate, change and load exchange information onto the information stockroom framework.
- Store and deal with the information in a multidimensional data set framework.
- Give information admittance to business experts and data innovation experts.
- Dissect the information by application programming.
- Present the information in a helpful configuration, like a diagram or table.

Various degrees of investigation are accessible:

- Artificial neural networks: Non-straight prescient models that learn through preparing and take after natural neural organizations in structure.
- Genetic algorithms: Optimization methods that utilize cycles like hereditary mix, transformation, and normal determination in a plan dependent on the ideas of common development.
- Decision trees: Tree-formed constructions that address sets of choices. These choices produce rules for the characterization of a dataset.
- Nearest neighbor method: A method that orders each record in a dataset dependent on a mix of the classes of the K-type record generally like it in a verifiable dataset. Some of the time called the k-closest neighbor procedure.
- Rule induction: The extraction of valuable assuming guidelines from information dependent on factual importance.

- Data visualization: The visual understanding of complex connections in multi-dimensional information. Designs devices are utilized to represent information connections.

2 Literature Review

- Current exercise in health facility is to acquire required data for diabetes analysis via a number of exams and terrific remedy is furnished based totally on diagnosis. Data mining algorithms performs a massive position in healthcare industries. Healthcare industries have giant extent databases. Using massive information analytics one can learn about large datasets and locate hidden information, hidden patterns to find out information from the statistics and predict effects accordingly. The work predicts diabetes by the use of exterior elements accountable for diabetes like Glucose, BMI, Age, Insulin, etc. It makes use of a pipeline model to raise classification accuracy [1].
- One of the most essential elements of the overall performance of recurrent neural community is the weight parameter. Its significance relies upon the initialization process. An appropriate determination of initialization technique will end resulting in a suitable overall performance of the network. There are many optimization strategies that are being used in the kingdom of the art. The most frequent optimization method is gradient descent. Many different optimization strategies have been developed based totally on the gradient fashion such as stochastic gradient decent, mini-batch gradient decent, Adagrad, Adam, and many others [2].
- Artificial neural networks represent allotted processing systems, comprised of densely interconnected, adaptive processing units, characterized with the aid of an inherent propensity for getting to know from trip and additionally discovering new knowledge. The notable functionality of self-learning and self-adapting of these mastering structures has installed them as effective equipment for sample cognizance and as integral aspect of many classification systems. Thus, they have been appreciably studied and broadly used in many purposes of synthetic brain [3].
- New advances in the laptop getting to know area, such as deep learning, have made the interaction between the optimization strategies and computing device mastering one of the most vital factors of superior computational science. Optimization strategies are proving to be quite essential in order to instruct fashions which are capable to extract statistics and patterns from massive volumes of data. Recent breakthroughs in illustration getting to know and deep getting to know exhibit that giant fashions dramatically enhance performance. As the quantity of mannequin parameters increase, basic implementations of optimization strategies on one single computing device are no longer feasible. Many purposes require fixing optimization troubles with a giant wide variety of parameters. Problems of

this scale are very frequent in the Big Data era. Therefore, it is vital to learn about the hassle of large-scale optimizations on disbursed systems [4].

- Machine learning (ML) techniques, over the ultimate few years, have been considered to exhibit an accelerated relevance to a range of objectives, which include hazard assessment. This wealthy information may additionally be beneficial for some decisive steps to represent ailment threat and progression. ML strategies appear to be an attractive choice for the prevention of T2DM [5].

Diabetes is a sickness the place where blood sugar stages or blood glucose are high. It has three kinds –Type 1, Type two and Gestational diabetes. Type I diabetes, the place the human physique cannot produce adequate insulin. Type two diabetes, is a frequent kind the place the human physique cannot create or make use of insulin well. Gestational diabetes will happen for the duration of pregnancy. Diabetes is the principal motive of risky fitness troubles like coronary heart strokes, eye problems, anxious gadget disturbance and kidney problems. According to survey reviews in 2019, via International Diabetes Federation (IDF), the humans who have diabetes is close to 463 million. Researchers predicted that the variety of diabetes sufferers may additionally amplify to 642 million, i.e., one in ten adults. In this circumvention, the early prediction of diabetes mellitus is important with environment friendly strategies to minimize the dying rate. Among the diabetes cases, greater number belongs to category two diabetes [6].

3 Methodology

3.1 GMM Classifier Algorithm

At its most straightforward, GMM is likewise a kind of grouping calculation. As its name suggests, each bunch is demonstrated by an alternate Gaussian dissemination. Every information point might have been produced by any of the disseminations with a relating likelihood. Basically, every appropriation has some obligation regarding producing a specific information point applying the GMM classifier has two stages. In the preparation stage, a GMM classifier is prepared for each class. The likelihood appropriation capacities are determined for each class. In the application stage, a class of the greatest likelihood thickness work is relegated to this class. The use of GMM is exceptionally wide. For example, in picture handling and PC vision, customary picture division models regularly allot to one pixel only one elite example. In fluffy or delicate division, any example can have certain "possession" over any single pixel. The given case is examples of Gaussian, fluffy division which normally brings about Gaussian combinations. Additionally, GMM is likewise utilized as a classifier to assist the specialist with anticipating illness.

3.2 Artificial Neural Network

It is genuinely said that the working of ANN takes its foundations from the neural organization dwelling in the human mind. ANN works on something alluded to as Hidden State. These secret states are like neurons. Every one of this secret state is a transient structure which has a probabilistic conduct. This network of a particularly covered up state go about as an extension between the information and the yield. The current work utilizes ANN to anticipate diabetes. The best outcome, 0.89, is acquired when the quantity of the secret layer is 2, and the quantity of each secret neuron is 5. The $f(x) =$ Sigmoid was picked as an organization work. At the outset, the boundaries were instated arbitrarily. Concerning the advancement strategies, the SGD (stochastic angle slip) was picked. For higher precision, the other differentiation tests were directed also. In any case, because of the impediment of the sum and accessibility of the preparation informational indexes, the scope of the quantity of the secret layer is set from 1 to 3, and the quantity of the neurons is from 5 to 10.

3.3 Extreme Learning Machine (ELM)

Extreme learning machines are feed-forward neural network for arrangement or relapse with a solitary layer of covered up hubs, where the loads associating contributions to cover up hubs are haphazardly doled out and never refreshed. The loads between covered up hubs and yields are learned in a solitary advance, which basically sums to learning a direct model [7]. As indicated by their makers, these models can create great speculation execution and learn a huge number of times quicker than networks prepared utilizing back propagation. The least complex ELM preparing calculation learns a model of the structure.

$$Y = w_2\,(w_1) \tag{1}$$

where w_1 is the matrix of input-to-hidden Layer weights, (Y is some activation function, and w_2 is the matrix of Hidden-to-output-Layer weights. Generally, the w_1 would be filled with Gaussian random noise and w_2 can be estimated by least -squares fit. The current work also uses ELM with 10 neurons and $f(x) =$ sigmoid in the hidden layer.

3.4 Logistic Regression and SVM

In this model the needy variable is downright, specifically paired ward variable that is, the place where it can take just two qualities, “0” and “1”, which address results like pass/fail, win/lose, alive/dead or sound/debilitated. Calculated relapse

is utilized in different fields, including AI, most clinical fields, and sociologies. Strategic relapse was utilized to foresee whether a patient experiences diabetes, in view of seven noticed attributes of the patient. The support vector machine (SVM) was first proposed by Vapnik, and SVM is a bunch of related managed learning strategies consistently utilized in clinical determination for grouping and relapse. SVM all the while limits the experimental characterization mistake and amplifies the mathematical edge. So SVM is called Maximum Margin Classifiers. SVM is an overall calculation dependent on ensured hazard limits of factual learning hypothesis, the alleged primary danger minimization standard. SVMs can productively perform nonlinear order utilizing what is known as the portion stunt, certainly planning their contributions to high-dimensional component spaces. The bit stunt permits developing the classifier without unequivocally realizing the component space. As of late, SVM has pulled in a serious level of revenue in the AI research local area. A few late investigations have detailed that the (support vector machines (SVM) by and large are fit for conveying better as far as characterization exactness than the other information order calculations. SVM is a method reasonable for double characterization undertakings, so we pick SVM to anticipate diabetes.

4 Results and Discussion

The test used to be carried out involves the usage of python platform. Python is a high-level, interpreted, interactive and object-oriented scripting language. Its important facts are, it is Easy-to-learn, Easy-to-read, Easy-to-maintain. A vast fashionable library—Python's bulk of the library is very transportable and cross-platform well matched on UNIX, Windows, and Macintosh. The dataset used right here used to be the Pima Indians diabetes database. The dataset consists of 768 entries having 9 features. The entries correspond to take a look at every patient. The 9 aspects are:

- Pregnancies—Number of times pregnant
- Glucose Plasma—glucose concentration a 2 h in an oral glucose tolerance test
- Blood Pressure—Diastolic blood pressure (mm Hg)
- Skin Thickness—Triceps skin fold thickness (mm)
- Insulin—2-h serum insulin (mu U/ml)
- BMI—Body mass index (weight in kg/(height in m)ˆ2)
- Diabetes Pedigree Function—Diabetes pedigree function
- Age—Age (years)
- Outcome—Class variable (0 or 1) 268 of 768 are 1, the others are 0

As per the experiments carried out in previous works ANN has the best performance. The performance data is as given below (Table 1):

The figure given below clearly shows that the ANN algorithm is most efficient in prediction of diabetics. The limitation of ANN is that the efficiency of the ANN-based method depends on the amount of data. Larger the data, higher the efficiency of the method (Fig. 1; Tables 2, 3, 4, 5).

Table 1 Result comparison in execution time for diabetes techniques

Method	GMM classifier	Artificial neural network (ANN)	Extreme learning machine ELM	Logistic regression	Support vector machine SVM
Result in (%)	0.8	0.9	0.88	0.66	0.75

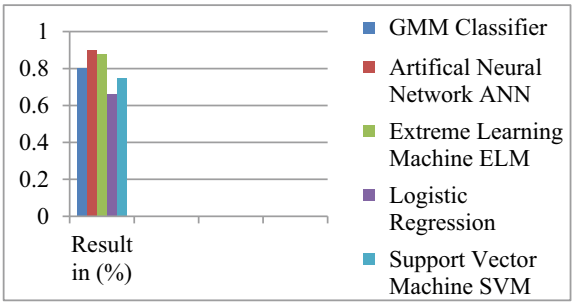


Fig. 1 Comparison of various dubieties prediction methods

Table 2 Result comparison between GMM and ANN

Method	GMM classifier	Artificial neural network (ANN)
Result in (%)	0.8	0.9

Table 3 Result comparison between ELM and ANN

Method	Artificial neural network (ANN)	Extreme learning machine (ELM)
Result in (%)	0.9	0.88

Table 4 Result comparison between ANN and logistic regression

Method	Artificial neural network (ANN)	Logistic regression
Result in (%)	0.9	0.66

Table 5 Result comparison between GMM and SVM

Method	Artificial neural network (ANN)	Support vector machine (SVM)
Result in (%)	0.9	0.75

As per the experimental results shown above following observations were found:-

- The overall performance of GMM classifier is closely established on range of components. When the variety of issue used to be one then the overall performance of the classifier was once very bad round 60%. With extend in factor count number

to two or higher the overall performance used to be high-quality which used to be round 80%. But as the quantity of elements was once multiplied the quantity of coaching facts additionally increased. The foremost drawback is if the mannequin is difficult the quantity of coaching statistics additionally wants to be large. Among the classifiers the GMM is the quickest algorithm, however overall performance is no longer consistent. This method suffers from the complexity of its O(NKD 3) approach, which makes it unusable for high-dimensional data.

- The overall performance of ANN used to be great though there used to be restrained availability of datasets. The quantity of neurons supplied used to be in vary 5 to 10 and wide variety of hidden layers furnished was once in vary 1 to three. The common accuracy reached used to be round 85–90% and used to be constant between the trade in wide variety of hidden layers and neurons. We got 2 layers to distribute the remaining n-1 neurons to the cells. The worst case would be that the nodes are equal distributed, which would result in O (nˆ2), where n is the no of neurons.
- ELM has validated true potentials to resolving regression and classification troubles and additionally gives a unified mastering platform with a considerable kind of characteristic mappings. Although accuracy of ELM used to be above 85% the important drawback is it is plenty quicker to train, however cannot encode greater than one layer of abstraction, so it cannot be "deep". Another trouble is though you can teach them definitely fast; you pay for it by means of having very gradual evaluation. For most applications, contrast pace is extra vital than coaching speed.
- Complexity of hidden layer output matrix is equal to O (NLD) and network output weight matrix calculation having time complexity equal to O ($L^3 + L^2 N + LNC$).
- The overall performance of the logistic regression algorithm was once the poorest even though it is simpler to implement, interpret, and very environment friendly to train. But as the complexity of the dataset will increase it has a tendency to over-fit. It is challenging to gain complicated relationships in the usage of logistic regression. More effective and compact algorithms such as neural networks can without problems outperform this algorithm. Its complexity for classification of the existing dataset is O((f + 1)cs).Where f is the no of features and c is no of data points.
- SVM algorithm is now not appropriate for massive facts sets. SVM does now not function very nicely when the information set has extra noise, i.e., goal instructions are overlapping. In instances the place the quantity of points for every information factor exceeds the variety of coaching records samples, the SVM will underperform. The quantity of elements used for the test was once 7. But as the range of facets will increase the algorithm may additionally now not operate as expected. For the present day test the accuracy was once round 70%. The time complexity for predicting can be given as O (kd) where k is no of support vectors and d is number of data points.

5 Conclusion

The survey specializes in conducting an observe to assess five one-of-a-kind records mining strategies. Those strategies are used to behavior the early prediction over diabetes, they are consisting of: GMM, ELM, SVM, Logistic regression and ANN. The cost of GMM is usually higher if the data set contains many distributions. ANN performance is affected by the larger number of distributions and the number of observed data points. As a result, smaller weights can improve the neural network's stability. An ELM is much easier to learn, but it can only encode one level of abstraction at a time, thus it cannot be "deep." For big data sets, the SVM algorithm is ineffective. When there is more noise in the data set, SVM does not perform well. Because logistic regression has a linear decision surface, it cannot tackle nonlinear issues. The both synthetic and real-global experiments display that ANN approach is most effective, more dependable and competitively sturdy close to numerous cluster sizes, shapes and densities. After a neural community is efficaciously skilled, its class accuracy is trusted architecture, but mostly on the values of its weights. Destiny work is focused on the implementation of a weight confined neural network schooling set of rules for green prediction of diabetes.

References

1. Alqushaibi A, Abdulkadir SJ, Rais H, Al-Tashi Q (2020) A review of weight optimization techniques in recurrent neural networks. In: 2020 international conference on computational intelligence (ICCI)
2. Mujumdar A, Vaidehi V (2019) Diabetes prediction using machine learning algorithms. In: International conference on recent trends in Advanced computing 2019. https://doi.org/10.1016/j.procs.2020.01.047
3. Kumar PBM, Perumal RS, Nadesh RK, Arivuselvan K, Type 2: diabetes mellitus prediction using deep neural networks classifier. https://doi.org/10.1016/j.ijcce.2020.10.002
4. Perveen S, Shahbaz M, Ansari MS, Keshavjee K, Guergachi A (2020) A hybrid approach for modelling type 2 diebeties mellitus progression. www.frontiersin.org, 10, Article 107, Jan (2020)
5. Najafabadi MM, Khoshgoftaar TM, Villanustre F, Holt J (2017) Large-scale distributed L-BFGS, Najafabadi et al. J Big Data 4:22. https://doi.org/10.1186/s40537-017-0084-5
6. Livieris IE (2018) Optimal algorithms and the BFGS updating techniques for solving unconstrained nonlinear minimization problems. Informatics 2018, xx, 1. https://doi.org/10.3390/informaticsxx010001
7. https://www.healthline.com/health/difference-between-type-1-and-type-2-diabetes
8. Shona D, Shobana A (2015) A survey on intrusion detection using data mining technique. Int J Innovative Res Comput Commun Eng 3(12), Dec 2015
9. Chourse S, Richhariya V (2014) Survey paper on intrusion detection using data mining techniques. Int J Emerg Technol Adv Eng 4(8), Aug 2014

Utilization of Intellectual Learning Methods in Forest Conservation: Smart Growth and Upcoming Challenges

R. Vasanth and A. Pandian

Abstract Nowadays, forest ecology is applied in various disciplines, mainly in machine learning, which is a crucial branch of artificial intelligence. Here, widely used machine learning approaches and its uses of forest natural science were reviewed over the last ten years. Granting ML techniques help classify, model, and predict forest ecology studies, the further development of ML technology is narrowed by the insufficient of applicable data and greater threshold applications. On the other hand, the combination and use of different algorithms, as well as enriched discussion and group effort flanked by environmental scientists and ML creators, still current significant experiments and challenges for upcoming environmental investigation. We trust in "big data," where environmentalists resolve soon have right to use to additional kinds of data such as sound and video, forthcoming uses of machine learning in environmental science will turn out to be an progressively smart device for environmentalists, potentially inaugurate a novel ways of investigation in forest environmental science.

Keywords Artificial intelligence · Classification · Prediction · Species dispersion system · Carbon cycles · Forest ecology

1 Introduction

Forests cowl nearly 30% of the arena's land-living location and are the leading native surroundings taking place on terrain. Such as, forest environments traditionally acknowledged tons responsiveness as of experts which has been looking en route for realize the compound interfaces in the middle of diverse environmental strategies that initiate the underlying forces of the structures [1]. The new boom in the ease of use of big quantities of statistics and the improvement of records evaluation approaches able to dealing with massive datasets are offering new possibilities to take a look at those complicated systems [2]. Gadget getting to know (ML) is a

R. Vasanth (✉) · A. Pandian
SRM Institute of Science and Technology, Kattankulathur 603203, India
e-mail: vr1116@srmist.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_28

vital department of artificial intelligence, which offers a few giant benefits in excess of conventional numerical techniques for examining woodland environmental facts when satisfactorily big datasets are obtainable as model schooling units. The ML utility tactics especially encompass

(i) Choosing significant information and its data processing
(ii) Choosing of satisfactory procedures
(iii) First-rate analysis answers.

Because the nineteen nineties, ML has progressively more been used in ecological disciplines absorbed on several fields of studies which encompassed marine science, geology, and meteorology, however, they hardly ever said that how ML changed into second-hand to examine wooded area ecologies; but, these techniques progressively emerge as widely held in woodland environment studies as reproduced by way of accumulating quantity of guides during the decagon [3]. On this studies area, ML methods supplied powerful and green methods to cope with information that have been indiscriminate, great proportion, and enclosed compound interfaces together with lacking values. Now, we gift a brief assessment of ML procedures carried out toward woodland surroundings examines together through an explanation of obstacles that save you comprehensive ML adoption. To accomplish that, we briefly:

(i) Refer to the over-all outline of ML after which awareness 3 unique ML algorithms for wooded area atmosphere studies.
(ii) Overview and synthesize currently put out packages of ML in woodland ecologies.
(iii) Deliberate dual blocks of ML in woodland environmental science and a few destiny applicable studies
(iv) Current key ends with view point at the use of ML techniques.

2 Machine Learning

2.1 Framework

ML generation enables computer systems discover designs to enhance prediction by information and makes use of these styles. The “education” and “forecasting” procedures in ML can match up to the hominid “simplification” and “assumption” procedures. Evenhanded as revel in critical in studying, ancient datasets perform a conclusive function in ML [4]. In the statistics analytics area, ML is a technique, and this is recycled to plan complicated fashions that enforce themselves for forecast. In eco-friendly or conservation research, these logical simulations permit investigators to “output reliable, repeatable consequences” and learn “unidentified interactions” thru studying from significant datasets [5].

2.2 Unique Three ML Procedures Used in Forest Environmental Science

ML may be divided into two major categories: supervised learning and unsupervised learning. Supervised learning presents a clean anticipation of outcome afterward enter examples consisting of class and regression have qualified thru the simulation [6]. Unsupervised learning is pretty random in what kind of outcome is produced afterward input examples had been educated to the model. A standard instance is clustering, that is, take along composed alike gadgets. We offer a brief explanation of the 3 famous and extensively used ML procedures in woodland environments: tree-based AI techniques, artificial neural network (ANN), and support vector machine (SVM) [7].

Tree-based AI techniques

Tree-based AI techniques are among the most usually utilized regulated learning strategies. They are developed by two elements; branches and hubs. Tree-based ML techniques are worked by recursively parting a preparation test, utilizing various highlights from a dataset at every hub that parts the information most viably. The parting depends on taking in straightforward choice standards construed from the preparation information [8]. For the most part, tree-based ML strategies are basic and instinctive; to foresee a class mark or worth, we start from the highest point of the tree or the root and utilizing branches, go to the hubs by looking at highlights based on which will give the best split. Tree-based strategies likewise utilize the mean for constant factors or mode for absolute factors when mentioning forecasts on preparing observable facts in the districts they have a place with [9]. The techniques are adaptable and can be applied to one or the other grouping or relapse issues. Arrangement and regression trees (CART) are a usually utilized term by Leo Breiman, alluding to the adaptability of the strategies in addressing both direct and non-straight prescient displaying issues [10].

ANN

Counterfeit neural networks can be best depicted as the organically motivated reproductions that are performed on the PC to do a specific explicit arrangement of errands like bunching, characterization, design acknowledgment, and so forth. All in all, artificial neural networks are a naturally propelled organization of neurons (which are fake in nature) arranged to play out a particular arrangement of undertakings [8].

Sensible neural networks may be first-class understood as weighted coordinated diagrams, in which the hubs are fashioned via the counterfeit neurons and the affiliation between the neuron yields and neuron facts assets may be addressed through the coordinated edges with loads. The artificial neural network receives the facts signal from the outer international as an example and picture as a vector. Those records assets are then numerically assigned via the documentations $x(n)$ for every n number of data resources [7].

Each one of the facts is then accelerated by using its referring to hundreds (those masses are the subtleties used by the counterfeit neural businesses to take care of a specific issue). In everyday phrases, these loads normally deal with the energy of the interconnection among neurons in the faux neural company [5]. Every one of the weighted facts assets are summarized inside the registering unit (one more counterfeit neuron). Inside the event that the weighted general likens to not anything, a bent is delivered to make the yield non-zero or more than probably proportional as much as the framework's response. To preserve the reaction within the constraints of the right worth, a specific restriction esteem is benchmarked. And in a while the quantity of weighted facts sources is long past thru the actuation work [11].

SVM

The SVM calculation was gotten from factual learning hypothesis. The calculation depends on the underlying danger minimization rule; it can pack the assortment of crude information to a help vector set (normally 3–5% of the crude information) and figure out how to acquire an arrangement choice capacity. The essential thought is to build a hyper plane as the choice surface so the stretch between the positive and negative mode is most extreme [7].

The empty circles and open squares address the two sorts of preparing tests. H is the grouping line that is accurately isolated. H1 and H2 not just pass through the focuses that are closest to the different example types yet in addition are corresponding to the arrangement line, and the distance between the two lines is known as the order stretch. As per the guideline of experimental danger minimization hypothesis, the genuine SVM hazard is chosen by the recipe $R(\omega) \leq \text{Remp}(\omega) + \Phi$, where $R(\omega)$ addresses the real danger, $\text{Remp}(\omega)$ addresses the exact danger, and Φ addresses the certainty stretch. Complete division guarantees $\text{Remp}(\omega) = 0$, and a greatest span guarantees the base scope of the certainty stretch Φ with the goal that the genuine danger is limited [10] (Tables 1 and 2).

3 Utilization of AI Procedures in Woodland Nature

ML has been generally received and incorporated by scientists considering expanding worries over timberland environments, including

(a) Species dispersion demonstrating.
(b) C cycles.
(c) Risk assessment and forecast.

In this article, we just spotlight on generally later (for the most part after 2008) distributed utilizations of ML in woods environments.

Table 1 Pros and cons of tree-based AI techniques, ANN, and SVM

Algorithms	Pros	Cons
Tree-based AI techniques	The nonlinear relationships between the parameters now do not affect the performance of the tree; Thus, tree technologies require relatively little customer-seeking in setting up facts. Ease of interpretation and knowledge are great advantages of using wood in analyzes. Tree techniques can also be used as a first-order extension of a huge database, even if its period is independent of the database period. CART is strong for the results of outliers within the output. RF can efficiently reduce the chance of overfitting	Missing information could have an impact on preference trees, and overfitting also can moreover result. It has more difficult to deal with missing information
ANN	ANNs have the capacity to concentrate further to display nonlinear and complex connections. They are likewise solid and flaw lenient to boisterous information. ANNs have a strong limit with respect to resemble preparing	Learning system cannot be determined during a black box, which finishes up in output, and this can be arduous to explain. ANNs are not ready to become alert to the relative significance and outcomes of character environmental variables
SVM	SVMs can display nonlinear decision limits, and there are different bits to decision out from. It is likewise very strong nearer to overfitting, explicitly in unbalanced dimensional space. SVMs can be proficient with a couple of significant pixels and can healthy controlled data	SVM is reminiscence intensive, harder to tune due to the importance of choosing the correct kernel, and now does not scale well to large datasets. Extrapolating from the bad version gives the bottom line when the above facts are inconsistent like the version. In all fairness, it mostly relies on the last facts as guide vectors

3.1 Species Dispersion System

Species conveyance displaying (SDM) is likewise known under different names including environment envelope-demonstrating, natural surroundings demonstrating, and (natural or biological) specialty demonstrating. The point of SDM is to appraise the likeness of the conditions at any site to the conditions at the areas of known event (and maybe of non-event) of a wonder. A typical utilization of this strategy is to foresee species ranges with environment information as indicators [6].

Table 2 The application and highlights of machine learning in forest ecology

Application	Methodology	Highlights
The predicted distribution of species with climate change	RF	Prediction accuracy can be improved by using RF
	SVM	When observing fewer records, SVM performs well
An analysis of the carbon and energy fluxes	ANN	In combination with traditional models, ANNs reduce uncertainty in predictions. ANNs are excellent at mining data
Assessing and predicting hazards	ANN	Data of high dimensions can be handled by ANNs, and they can be applied effectively. It is possible for ANNs to map nonlinearly
	SVM	In terms of resolving classification issues, SVMs are an excellent tool
Timberland management	ANN	ANNs are acceptable at anticipating over-the-ground biomass, wood volume, tree tallness, and stem tighten
	RF	Joining RF and other spatial addition approaches have extraordinary potential

In SDM, the accompanying advances are generally taken: (1) areas of event of an animal varieties (or other marvel) are incorporated; (2) estimations of ecological indicator factors (like environment) at these areas are separated from spatial datasets; (3) the natural qualities are utilized to fit a model to gage comparability to the locales of event, or another measure like wealth of the species; (4) the model is utilized to anticipate the variable of premium across the district of revenue (and may be for a future or past environment).

3.2 Carbon Cycles

Conventional demonstrating approaches have an extraordinary ability to evaluate and foresee C cycles, which for the most part relies upon the information utilized for definition and recognizable proof of information yield connections and can be up scaled from neighborhood to local or worldwide scales. Be that as it may, the versatility of these models is regularly unacceptable, which for the most part prompts dubious expectations if spatial and transient conditions change. Luckily, the flexibility of ANNs to natural conditions is solid if preparing and test information are adequate. Therefore, in 2016, they led another examination utilizing 11 ML calculations while applying four wide methodologies (tree-based techniques, relapse splines, neural organizations, and portion strategies) to foresee CO_2 and energy transition across different biological system types. In their examination, better forecasts of motion were accomplished for forested and calm districts contrasted and zones under outrageous environment conditions or with less information. They found that ML strategies had the option to foresee across-site inconstancy and mean occasional patterns of the

noticed transition well (R2 > 0.7); nonetheless, they acquired vulnerability results with 8-day deviations from the mean occasional cycle (R2 < 0.5) [10].

3.3 Risk Assessment and Forecast

Regular dangers brought about by bug episodes are among the most far and wide unsettling influences that sway the equilibrium of woodland environments in various areas. A directed characterization innovation joined with an improved component determination way to deal with evaluate the capability of hyper spectral symbolism, and they created a guide of bark-insect initiated tree mortality. Their outcomes appeared that the general exactness (OA) of planning dead trees was 84–96%, and the OA of the partition among sound and dead trees was 94–97%. Spatial-reliance examination, conventional kriging, and neural organization based relapse displaying to research the examples of bark-scarab flare-ups and the easygoing connections in auxiliary Norway tidy woods biological systems. They deduced that two bark-scarab episodes (1995–1999 and 2001–2004) brought about impractical optional tidy woods in Central Europe [5].

ML is an amazing grouping, displaying, and forecast apparatus in timberland environment research. In particular, ML models have a higher exactness and quicker limit in settling complex issues, breaking down connections, and anticipating nonlinear framework conduct. This strength of ML techniques, be that as it may, requires more thorough preparing and test information. Nonetheless, long haul and profoundly exact observing are costly; furthermore, information assortment, stockpiling, and refreshing can be upset by diminished subsidizing, instrument disappointment, restrictions of recorded advancements, obstruction by human exercises, etc. Albeit advantageous instruments (e.g., MATLAB and the R programming language) have effectively given amazing and agreeable UIs that intend to diminish numerous client hindrances, ML clients actually need to dominate some vital abilities that numerous biologists tragically need to investigate boundaries. At last, if ML is to be all the more every now and again utilized in woodland biology, scientists need better numerical capability and additional preparation abilities in programming to guarantee that they get calculations and potential issues, for example, over fitting [7].

4 Summary

This revision principally shares out with prose connected to ML applications attentive on forestry natural science. On the other hand, in spite of our reinforcement for better usage of ML, ML methods are not intended toward and will not ever be able to response all questions associated to woodland ecology s. ML approaches, however, be responsible for numerous valuable tools that should be more analytically considered to contract with a number of relevant environmental difficulties. In this study,

we proposed one direction for future research on this subject. The strength of ML techniques has always been data mining, particularly deep mining. ML techniques will be the ideal alternative for ecologists in the face of an influx of a large amount of research data, especially on a worldwide scale, as more and more data sharing becomes a reality.

References

1. Patel HH, Prajapati P (2018) Study and analysis of decision tree based classification algorithms. Int J Comput Sci Eng 6(10):74–78
2. Navada A, Ansari AN, Patil S, Sonkamble BA (2011). Overview of use of decision tree algorithms in machine learning. In: IEEE control and system graduate research colloquium, 978-1-4577-0339
3. Jijo BT, Abdulazeez AM (2021) Classification based on decision tree algorithm for machine learning. J Appl Sci Technol Trends 02(01):20–28
4. Patel BR, Rana KK (2014) A survey on decision tree algorithm for classification. Int J Eng Dev Res 2(1). ISSN: 2321-9939
5. Zhang Y, Zhu Y, Lin S, Liu X (2011) Application of least squares support vector machine in fault diagnosis. In: Liu C, Chang J, Yang A (eds) ICICA 2011, Part II, vol 244. CCIS. Springer, Heidelberg, pp 192–200
6. Fabian P, Stąpor K (2017) Developing a new SVM classifier for the extended ES protein structure prediction. In: Federated conference on computer science and information systems, Prague, pp 169–172
7. Agarwal DK, Kumar R (2016) Spam filtering using SVM with different kernel functions. Int J Computer Appl 136(5), Feb 2016
8. Astuti W, Akmeliawati R, Sediono W, Salami MJE (2014) Hybrid technique using singular value decomposition (SVD) and support vector machine (SVM) approach for earthquake prediction. IEEE J Sel Top Appl Earth Obs Remote Sens 7(5):1719–1728
9. Zhang Y, Phillips P, Wang S, Ji G, Yang J, Wu J (2016) Fruit classification by biogeography-based optimization and feedforward neural network. Expert Syst 33(3):239–253
10. Karaboga D, Ozturk C (2009) Neural networks training by artificial bee colony algorithm on pattern classification. Neural Netw World 19(3):279
11. Garro BA, Rodríguez K, Vázquez RA (2016) Classification of DNA microarrays using artificial neural networks and ABC algorithm. Appl Soft Comput 38:548–560

A Review on Deep Learning Techniques for Saliency Detection

Kokila Paramanandam and R. Kanagavalli

Abstract Salient object detection (SOD) has an important role in computer vision and digital image processing especially in the field of medical, ecology and transportation etc. At the same time a lot of complexity exists for detection of salient objects due to the effect of light, weather and density of the image. This review paper is proposed to summarize the existing implementation and recent technology development in the SOD. The major attention is given on reviewing deep learning techniques and edge detection techniques for SOD. From this work it is observed that the use of deep learning along with convolution neural network detects the objects accurately in less time, and by reviewing different edge detection method it is noticed that by incorporating edge detection methods can detect the objects efficiently even if it is present at clumsy and occluded regions.

Keywords Salient object detection · Saliency map · Edge detection · Deep learning · CNN

1 Introduction

Object detection has wide range of application in the field like robotics, autonomous driving, security etc., and because of these applications a lot of new inventions are arising day by day. Even though the requirement is more, identifying salient object is a challenging task in domain of digital image processing and computer vision because of its appearance change, object pattern, illumination deformation, blur, noise and occlusion etc. This paper concentrates on providing in depth details of several techniques that have been used in area of image processing for detecting

K. Paramanandam (✉)
Computer Science & Engineering, PES University, Electronic City, Bengaluru, Karnataka 560100, India
e-mail: kokila.p.anand@gmail.com

R. Kanagavalli
Information Science and Engineering, The Oxford College of Engineering, Bengaluru, Karnataka 560 068, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_29

salient object efficiently by overcoming the above mentioned challenges and also represented many state of art techniques that exists for saliency mapping. Salient object detection based on CNN technique permits a comfort object representation with a detailed learning ability of an object with less time consumption. Due to that advantage, deep learning approaches have been proposed in literature surveys [1]. Survey on edge detection techniques on saliency detection has proceeded to let us know how important it is to preserve the edges of an object to attain accuracy. Since use of CNN makes the smoothing edge at some of the cases, incorporating edge detection methods could detect the image which has multilayer features more accurately [2].

2 Review on Salient Object Detection Techniques

Sun et al. [3] Presented a MAP method for detecting salient object. Using MAP they calculated the preliminary saliency map of all noticeable objects in a picture and the objects that are partially occluded are used as background for further process. Authors mentioned that all existing techniques assume image boundaries as background but for simplicity in their proposed model only occluded objects are considered as background and because of that, this technique has better efficiency when tested with standard datasets. Figure 1 shows the performance of MAP on salient object detection.

To identify the objects from complex scene [4] developed an innovative frame work based on the concept of human visual cortex pattern that uses colour sensitivity and brightness for detecting the objects boundaries in an image. To supress the unwanted edges in the texture element, spatial sparseness constraint was used for neural response. From the experimental results, the developed model has better flexibility on capturing object on grayscale images and also in colour images. Detecting

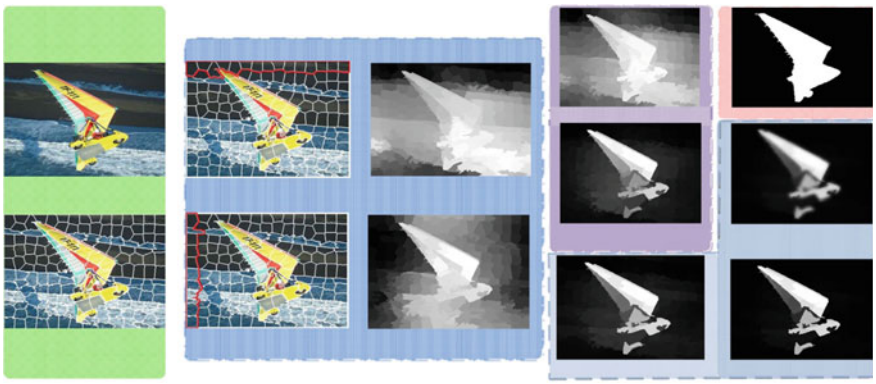


Fig. 1 Output of MAP on SOD [3]

object from supressed region is complex task, [5] introduced an approach that utilizes region cues and compactness for SOD. By combining the mentioned two features it generates pixel values uniformly on all salient objects which were used for better detection. While experimented with four benchmark datasets this approach attained good recall and F-measure than existing methods. Fu et al. [6] developed Ncut model that has good accuracy for object detection in over segmented regions as well as for complex background. It considers intrinsic colour and edge information to detect an object in a complex and clustered background. The model works on ordering the saliency value based on energy minimization rate to obtain clustered information in the scene. Model was experimented in four benchmark datasets and attained good performance on SOD in a clustered background. A framework was proposed by Yang et al. [7] for SOD in chaotic environment. This works in the form of two-pathway searching fashion i.e. during first parse contour based spatial prior technique is used to roughly identify an object and its location in an image. And during second parse local features are considered to estimate weight of the pixels and predict the exact location of salient object. When this framework experimented with six large datasets it acquired good efficiency irrespective to the size of object in the image. Li et al. [8] introduced diffusion based saliency region detection method which works on two layers. Initially a graph is created by connecting points to its neighbours that shares common boundaries. Then by using spatial variance, super weighted pixels are considered as seed vectors and it compares the seed vector with graph value to make sure the presence of saliency. This method clearly distinguishes the difference between the saliency seed from background seed. Moreover it was experimented on five benchmark datasets and obtained good performance in various evaluation metrics. Fig. 2 shows the working model of proposed system using seed values.

Fu et al. [9], Wang and Shen [10] provided a novel framework called RegionNet, which had been developed based on fast region CNN architecture for detecting objects. This framework overcomes the limitation of existing methods to identify

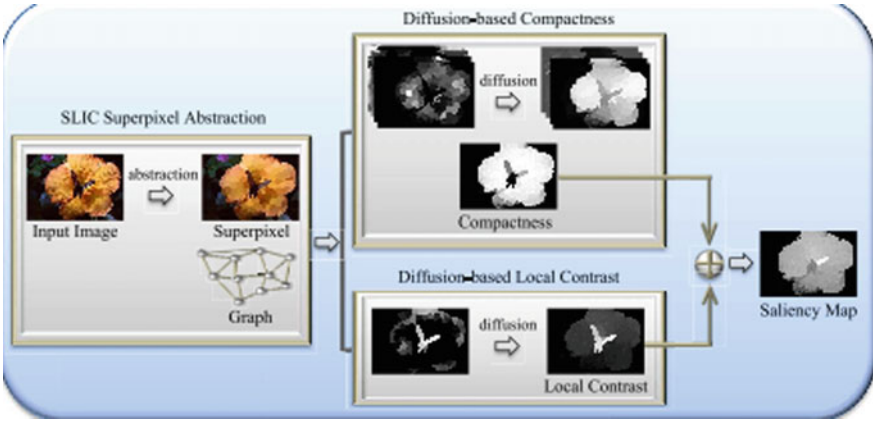


Fig. 2 Working model of [8]

objects in blurry boundaries and with less context information. RegionNet uses multiscale context to relate saliency regions to global scenes and to preserve the sharp edges of object boundaries. This framework achieves optimized performance when experimented with six RGB benchmark datasets. Kan et al. [11] proposed a multipass model for saliency mapping based on pixel classification technique. This model predicts pixel values of objects and builds a bounding box in an image. Within the constructed bounding box pixels are classified by mean of texture and convolution features. Through multipass on the different bounding box on the image the proposed work was able to identify the salient objects accurately. Han et al. [12] presented a hypergraph on adaptive multi scale (HAM) to determine the salient object. All colour images have pixel values in RGB range, so here multiple hypergraphs were constructed based on the adaptive scale and finally multiple adaptive scale hypergraphs are clustered together using agglomerative mean shift method. It is summarized that the experiment on HAM has shown improved performance on SOD particularly the images with least pixel values. Tang et al. [13] proposed a framework to correct false salient detection regions using boundary optimization method. The framework resolves false detection by latent segmentation method. The latent segmentation is used to fill the gap between high valued saliency maps lying inside and outside of the estimated object boundary. Which means latent segmentation method clearly differentiates salient object with the background. It is concluded; this framework improved its performance when tested with three state-of-art datasets. Figure 3 shows the output of boundary optimization framework.

Yuan et al. [14] proposed a framework built upon deep neural network for salient mapping. The proposed framework constructed adepth convolution network which uses dense value and sparse value for object prediction. It uses low level features and contours to estimate the position of the object. Furthermore convolution network conducts the saliency classification by considering the superpixels generated by dense and sparse values. The end result of this framework had good performance when experimented with six differently available open source datasets. Songtao et al. [15] defined a model considering region covariance and the global features. To construct the local salient map, the region covariance with various scale space is extracted from

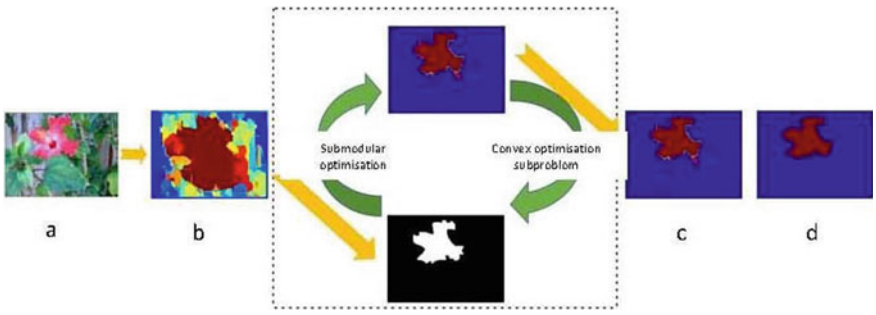


Fig. 3 Workflow of boundary optimization framework with final result [13]

the image region, and these scale values are converted into sigma features. On the other side gray scale and density estimation is considered to generate a global salient map. By combining the local and the global salient values saliency detection is done in an accurate manner. Concluded result of this model shown better performance on qualitative wise and quantitative wise. A multiscale deep-neural-network framework was suggested by Xiao et al. [16], Han et al. [12], Xiao et al. [17] for SOD. These frameworks use low-level features and high-level features to effectively identify the salient objects in sharp boundaries. Along with it deconvolutional network is used to get the multiscale feature to get the series of saliency map. By fusing both convolutional and deconvolutional layers the super pixel values in an image are calculated to frame the saliency map. Final result shows the proposed model significantly outperforms on six SOD benchmark datasets.

Zhugeet et al. [18] introduced FCN framework that utilizes boundary values and multilevel features values to detect an object. The estimation of saliency is done by FCN by extracting the features and aggregating it with same resolution levels. The boundary feature of the object is extracted by boundary information extraction method. Finally, the attention based fusion module fuses boundary information into salient regions to get accurate results. It is concluded that this framework has better results when experimented and analysed with four benchmark datasets. Cholakkal et al. [19] developed a supervised top down (TD) saliency framework for SOD that uses only the binary labels to point the presence or absence of saliency in a picture. The TD saliency region is generated by probabilistic contribution of the image which is computed by backtracking strategy. Bottom up (BU) saliency approach is used to analyse the finest saliency maps. The group of saliency maps formed by BU and TD are combined together and produce superpixel values and it is used to train the classifier for correct object prediction. Performance of proposed framework has quantitatively better results when experimented with seven challenging datasets. Guanbin and Yizhou [20] developed a model which predicts salient object in a single pass. This method overcomes time consuming problems that are faced by existing methods. The proposed method has two components for dense prediction and for sparse inference of saliency. FCN is used for dense estimation and segment level spatial method is used for sparse inference of saliency. Fusing both the methods determines the weight of saliency map and author concluded as this model had good performance in terms of many evaluation metrics. Manno-Kovacs [21] designed an active contour model to extract the salient object, this method deals with the object that resides in complex contour and cluttered background and identifies it with high accuracy. By combining the shape contour and texture feature for discovering multiple objects is also possible. Evaluation method of this model has higher efficiency than the existing methods.

3 Deep Learning in Salient Object Detection

Object detection is finding and classifying various numbers of objects in an image through localization, classification and segmentation methods. Localization is used

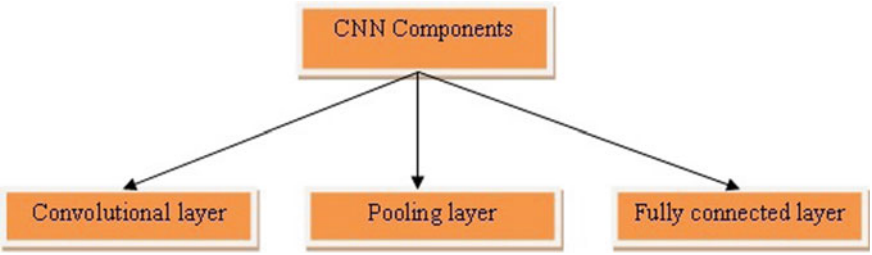


Figure 4. Components of CNN [22]

to find the location of an object, classification is to classify the object with its features and finally segmentation is used to find the pixel values of the objects for prediction. Moreover detecting multiple objects in an image is complex work because the presence objects are not known earlier. Therefore deep learning models are used to attain the goal. Deep learning model is in advance on detecting objects to predict correct number of objects in fast efficient way by mean of convolutional neural network, here the window of fixed size slides over the image for predicting the objects. Fig. 4 shows the components of CNN.

Wang et al. [23] proposes two CNNs to attainregion information and collective contrast in an image. Saliency map is predicted by fusing both deep CNNs [24, 25]. Whereas [26] mentioned a novel superpixel wise CNN method for better representation of saliency. Here saliency detection is done by binary labelling pattern. Liu et al. [27] provided few types of CNN based object detection architecture to get on go in this domain.

3.1 Region with CNN

Pierre et al. [28] proposed a multi-sale sliding window algorithm using CNN for object detection. Normally in this method an image is divided into several region and then it is classified, hence because of that it has high computation time. Soon after this, Region with CNN(R-CNN) by Girshick et al. [29] has shown 50% improvement on object detection from previous methods. Here selective search is used to extracts objects from each area using CNN and classifies that region using SVM. The entire work of identifying on object works based on varying scales, colors, textures, and enclosure. R-CNN model is slow on prediction; it takes nearly 50 s to predict an object in an image. Fig. 5 shows R-CNN architecture.

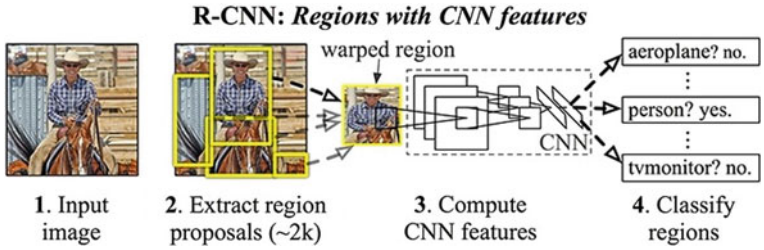


Fig. 5 R-CNN architecture [29]

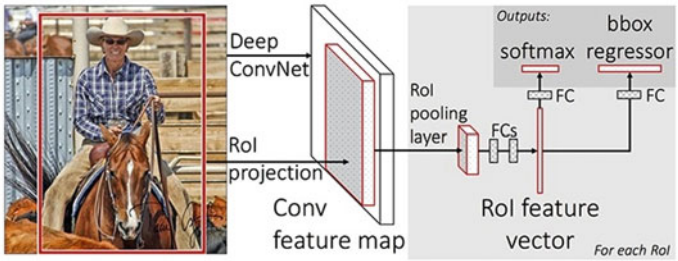


Fig. 6. Fast region with CNN architecture [30]

3.2 Fast Region with CNN

Fast region with CNN was introduced by Girshick et al. [30] it had better efficiency by applying selective search to entire image and generating region of interest (ROI). ROI pooling is applied to all area in image to predict the objects. Fast R-CNN has better performance (two seconds to predict an object) than R-CNN but still considered to be slower for large datasets. Figure 6 shows Fast R-CNN architecture.

3.3 Faster Region with CNN

Faster R-CNN [31] has overcome many of the limitation of Fast RCNN. It uses RPN to detect the image by mining the features of it. This method takes less than a second for object detection but there exist limitation like it requires multiple passes for identifying all objects in a scene. Fig. 7 shows Faster Region with CNN architecture.

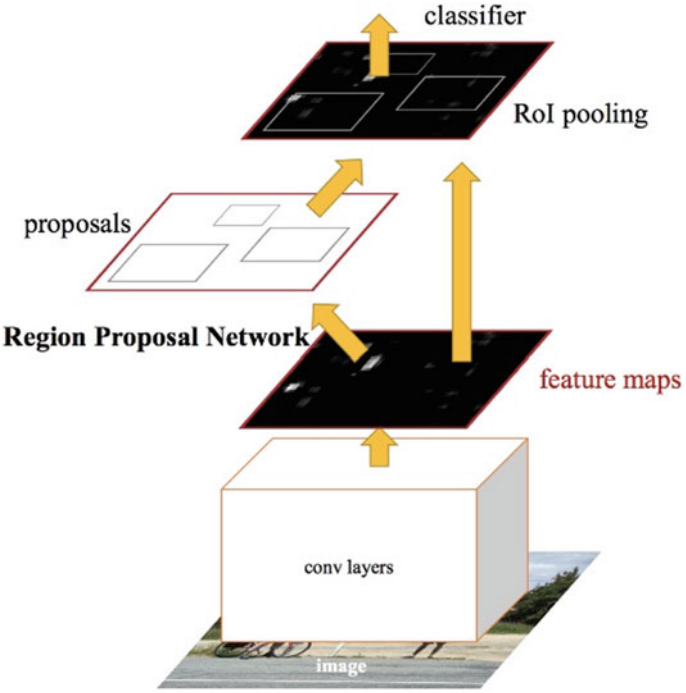


Fig. 7 Faster R-CNN architecture [32]

3.4 Yolo

Redmon et al. [33] proposed YOLO is an unique algorithm for object detection. YOLO is completely different from previously mentioned region based method because of its concept of dividing whole images into many grids and to consider it as bounding box. An assured pixel value is computed from each box and they are clustered together for identifying objects. This method has more accuracy but fails to detect small objects in image. Fig. 8 shows the working model of YOLO.

3.5 SSD

Liu et al. [27] developed SSD which overcomes the drawback of YOLO by using multiple sized convolutional feature maps and region based FCN for good performance and speed. Unlike YOLO, SSD usesgroup of bounding box with different scale for detecting the object of various sizes in an image. Figure 9 shows the difference between SSD layouts with respect to YOLO architecture.

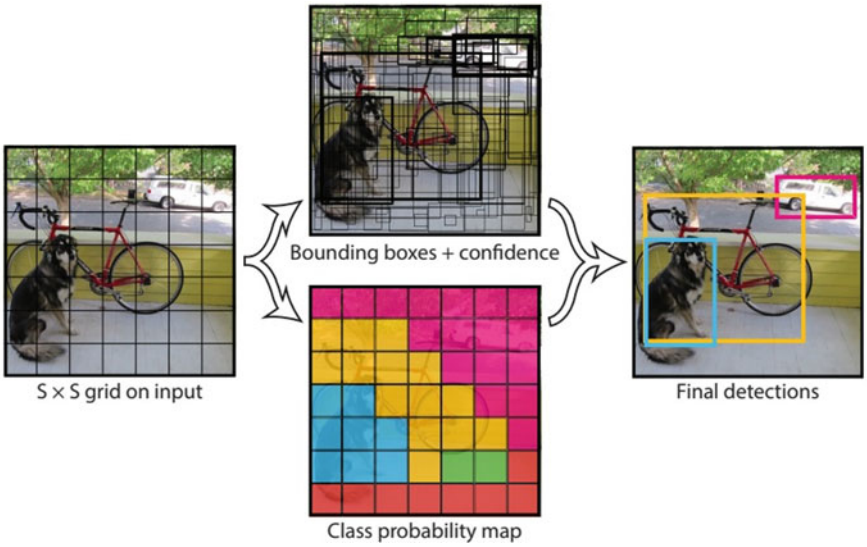


Fig. 8 YOLO Model [33]

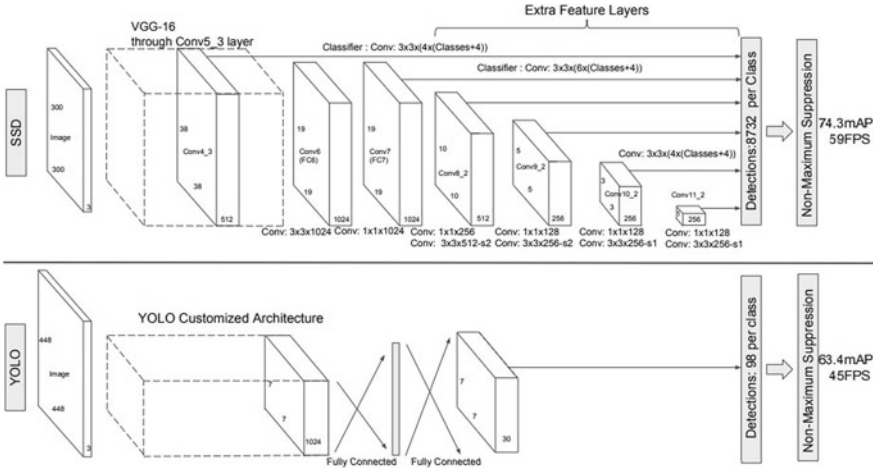


Fig. 9 SSD layoutvs. YOLO architecture [27]

4 Applications of Object Detection System

SOD plays an important role in Automated CCTV, Automatic image annotation, Automatic target recognition, Object counting, Digital watermarking, Medical imaging, Optical character recognition, Self-driving cars, Tracking objects, Ball

tracking in sports, Object recognition as image search, Manufacturing industry, Robotics and etc.

5 Conclusion

An in depth review of various researches for detecting salient objects is presented in this paper. The techniques available for SOD based on parameters like global features, local features, high-level, low-level features and multi-level features survey is done. This paper especially concentrated on the mechanism of SOD using deep learning techniques with respect to the challenges and its applications in real time. And this would help the researchers to identify the research gaps and opportunities in the computer vision and image processing domain.

References

1. Kawulok M, Benecki P, Piechaczek S, Hrynczenko K, Kostrzewa D, Nalepa J (2020) Deep learning for multiple-image super-resolution. IEEE Geosci Remote Sens Lett 17(6):1062–1066
2. Ye C, Slavakis K, Nakuci J, Muldoon SF, Medaglia J (2021) Fast sequential clustering in riemannian manifolds for dynamic and time-series-annotated multilayer networks. IEEE Open J Signal Proc
3. Sun J, Lu H, Liu X (2015) Saliency region detection based on markov absorption probabilities. IEEE Trans Image Process 24(5):1639–1649
4. Yang K, Gao S, Guo C, Li C, Li Y (2015) Boundary detection using double-opponency and spatial sparseness constraint. IEEE Trans Image Proc 24(8):2565–2578
5. Zhou L, Yang Z, Yuan Q, Zhou Z, Hu D (2015) Salient region detection via integrating diffusion-based compactness and local contrast. IEEE Trans Image Process 24(11):3308–3320
6. Fu K, Gong C, Gu IY, Yang J (2015) Normalized cut-based saliency detection by adaptive multi-level region merging. IEEE Trans Image Process 24(12):5671–5683
7. Yang K, Li H, Li C, Li Y (2016) A unified framework for salient structure detection by contour-guided visual search. IEEE Trans Image Process 25(8):3475–3488
8. Zhou L, Yang Z, Zhou Z, Hu D (2017) Salient region detection using diffusion process on a two-layer sparse graph. IEEE Trans Image Process 26(12):5882–5894
9. Fu K, Zhao Q, Gu IY (2019) Refinet: A deep segmentation assisted refinement network for salient object detection. IEEE Trans Multimedia 21(2):457–469
10. Wang W, Shen J (2018) Deep visual attention prediction. IEEE Trans Image Process 27(5):2368–2378
11. Wu K, Yu Y (2018) Automatic object extraction from images using deep neural networks and the level-set method. IET Image Process. ISSN 1751-9659
12. Han J, Chen H, Liu N, Yan C, Li X (2018) CNNs-based RGB-D saliency detection via cross-view transfer and multiview fusion. IEEE Trans Cybern 48(11):3171–3183
13. Tang L, Li H, Wu Q, Ngan KN (2018) Boundary-guided optimization framework for saliency refinement. IEEE Signal Process Lett 25(4):491–495
14. Yuan Y, Li C, Kim J, Cai W, Feng DD (2018) Dense and sparse labelling with multidimensional features for saliency detection. IEEE Trans Circuits Systems Video Technol 28(5):1130–1143
15. Songtao L, Ning J, Zhenxing L, Kanghui J (2018) Saliency detection of infrared image based on region covariance and global feature. J Syst Eng Electron 29(3):483–490

16. Xiao F, Deng W, Peng L, Cao C, Hu K, Gao X (2018) Multi-scale deep neural network for salient object detection. IET Image Process. 12(11):2036–2041
17. Xiao H, Feng J, Wei Y, Zhang M, Yan SS (2018) Deep salient object detection with dense connections and distraction diagnosis. IEEE Trans Multimedia, 20(12):3239–3251
18. Zhuge Y, Yang G, Zhang P, Lu H (2018) Boundary-guided feature aggregation network for salient object detection. IEEE Signal Proc Lett 25(12):1800–1804
19. Cholakkal H, Johnson J, Rajan D (2018) Backtracking spatial pyramid pooling-based image classifier for weakly supervised top-down salient object detection. IEEE Trans Image Process 27(12):6064–6078
20. Guanbin L, Yizhou Y (2018) Contrast-oriented deep neural networks for salient object detection. IEEE Trans Neural Netw Learning Syst 29(12):6038–6051
21. Manno-Kovacs (2019) Direction selective contour detection for salient objects. IEEE Trans Circuits Syst Video Technol 29(2):375–389
22. Gousia H, Shaima Q (2020) Optimization and acceleration of convolutional neural networks: a survey. J King Saud Univ Comput Inf Sci. ISSN 1319-1578. https://doi.org/10.1016/j.jksuci.2020.10.004
23. Wang Z, Liu Q, Ikenaga T (2017) Visual salience and stack extension based ghost removal for high-dynamic-range imaging. In: 2017 IEEE international conference on image processing (ICIP). Beijing, China, pp 2244–2248
24. Dong C, Loy CC, He K, Tang X (2016) Image super-resolution using deep convolutional networks. IEEE Trans Pattern Anal Mach Intell 38(2):295–307
25. Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans Image Process 26(7):3142–3155
26. He K, Gkioxari G, Dollr P, Girshick R (2017) Mask r-cnn. In: 2017 IEEE international conference on computer vision (ICCV), pp 2980–2988
27. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC (2016) SSD: Single shot MultiBox detector. Computer Vision—ECCV 2016, Lecture Notes in Computer Science, vol. 9905. Springer International Publishing, pp. 21–37. https://doi.org/10.1007/978-3-319-46448-0_2
28. Pierre S, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y (2014) OverFeat: integrated recognition, localization and detection using convolutional networks. Comput Vis Pattern Recogn, arXiv:1312.6229v4
29. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 IEEE conference on computer vision and pattern recognition, pp 580–587
30. Girshick R (2015) Fast r-cnn. (2015). In: IEEE international conference on computer vision (ICCV), pp 1440–1448
31. Wang Y, Su H, Zhang B, Hu X (2020) Learning reliable visual saliency for model explanations. IEEE Trans Multimedia 22(7):1796–1807
32. Ren S, He K, Girshick R, Sun J (2017) Faster r-cnn: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39:1137–1149
33. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: Unified, real-time object detection. Comput Vis Pattern Recognition. arXiv:1506.02640

Energy and Memory-Efficient Routing (E&MERP) for Wireless Sensor Networks

Karuppiah Lakshmi Prabha and Chenguttuvan Ezhilazhagan

Abstract Wireless sensor networks (WSNs) comprised of sensor nodes use the capabilities of sensing, computing, then communicating. But, sensors have limited energy and memory capacity to perform the above operations. Therefore, using the available energy and memory efficiently is perplexing problems in WSNs. Clusters-based routing protocols are situated to exploit network lifetime. In this work, we recommend an advanced combination of clustering and replacement algorithm to decrease energy consumption, prolong network lifetime and improve throughput during frequent transmission. The recommended protocol reduces energy depletion by finding the cluster head (CH) node with maximum residual energy, thereby preventing energy hole, thus, improving network lifetime, and manages memory by using replacement algorithm in case of buffer overflow. The simulation shows that the recommended protocol has extended network lifetime, increased data delivery than protocols like LEACH, LEACH-C, SEP and EEEHR protocols.

Keywords Memory management · Energy efficiency · Cluster · Network lifetime

1 Introduction

Wireless sensors are designed for sensing, wirelessly communicating and processing of information. The senor nodes are less expensive and power efficient, features in forming a network, which is used to monitor the target region. Collecting information on various parameters (e.g. temperature, humidity, etc.) for which the sensor nodes are deployed for monitoring, the sink node will receive and process the information and convey it to the user. Application of WSNs is wide range; includes surveillance of military areas, predicting disaster and environment monitoring and providing lot

K. Lakshmi Prabha (✉)
PSG Polytechnic College, Coimbatore, India
e-mail: klp.dec@psgpolytech.ac.in

C. Ezhilazhagan
Dr. NGP Institute of Technology, Coimbatore, India
e-mail: ezhilazhagan@drngpit.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_30

of contribution from the researchers to enhance the applications in military, industry and academic fields. Resource of the sensor node is very restricted in terms of ability to process, bandwidth, memory space and lifetime, and all these factors differentiate WSNs from the traditional ad hoc networks.

In many of the applications, all sensor nodes are generally driven by a battery, which is expected to operate long without recharging. Hence, the limited resources such as energy and memory are to be used efficiently, particularly when sensors are tightly deployed or frequent transmission occurs. Due to the high concentration of nodes, many nodes may transmit or receive data, lead to increased energy consumption and high memory storage which may result in decreased network lifetime and throughput. Data identifying, data handling and data transmission/reception are the major three parts of energy consumption in a sensor node; most serious way of energy consumption is transmitting/receiving information. Communication to the right next node is taken care by the CH, and hence, the choice of CH based on residual energy is very important. Clustering which uses hierarchical mechanism majorly reduces the latency and also increases the scalability of the network which has been comprehensively projected in previous works.

The remainder of this paper is organized as follows: Sect. 2 discusses the work related to proposed protocol, and Sect. 3 presents the proposed model. Section 4 provides the result of E&MERP efficacy and performance by means of simulations and an assessment made with LEACH, LEACH-C and SEP. Section 5 outlines the conclusion and future work.

2 Related Works

Low-cost WSNs are made with huge quantity of sensor nodes with restricted energy, storage, transmission and processing facility [1]. All the physical and environmental characteristics like moisture, pressure, temperature, combustion, movement were monitored by these sensor nodes of WSNs which are used for different real-time applications like surveillance devices, monitoring weather, tracking of objects, healthcare management, disaster assistance, mapping of biodiversity, building rational behaviour, precautionary management, etc.

In unmanned, unattended and aggressive environment [2] like Warfield, nuclear plant, chemical industries, refineries, jungle field were monitored with the deployment of sensor nodes. In such scenarios, replacing the battery becomes a tough job. Base station (sink) has affluent resources [3] with infinite storage, power and transmission capability which receives data from the sensor nodes of hard line environment.

Increasing the network lifetime and throughput of the network by using different routing protocols designed for WSNs [4–6]. Reducing the number of bits sending through the communication module improves the network lifetime by scheduling the communication. The energy hole in WSN prevents communicating to the sink though

the nearby nodes are capable of transmission. Energy hole problem was addressed by increasing number of sinks, and mobility of sink provides solution [7].

Heinzelman et al. [4] propose an energy-efficient routing protocol with centralized clustering control called low-energy adaptive clustering hierarchy (LEACH-C), and it is a modification of LEACH protocol for WSNs. Put together on the collected data and remaining energy about all nodes in the network, the CH was centrally placed and the cluster was configured with other nodes in the setup phase of LEACH-C protocol. Lindsey and Raghavendra [8] proposed an algorithm related to LEACH, called power-efficient gathering in sensor information systems (PEGASIS). The energy utilized is very high for receiving and sending circuits compared to the amplifying circuits for a sensor node for particular range of distance. Stable election protocol (SEP) [9] was designed to extend the time interval of the first node which dies in heterogeneous network that is essential in numerous uses where the reliable feedback provided by the sensor network.

Another reason for increased network lifetime is unequal cluster size model [10] in WSNs, leading to uniform energy dissipation among cluster head nodes. Energy-efficient clustering scheme (EECS) [11] proposes cluster head distribution without any iteration, but the weighted function of a general node was brought to make decision, so that suitable cluster can be formed for WSNs with mobile sink.

The CH is selected based on the battery voltage level as addressed in fail safe fault tolerant algorithm [12, 13]. The battery potential level is the vital parameter to calculate the node's lasting energy. The energy used up by the node depends on quantity of bits transmitted and the distance to the receiver [6, 14, 15].

The CH selection impacts throughput of WSNs and also the network lifetime. Energy-efficient energy hole repelling (EEEHR) [16] algorithm developed for delay tolerant wireless sensor network with static nodes takes care of hotspot, and energy hole in the network makes the hole still larger when extra load to the node is provided. The problem is solved by increasing clusters and reducing the CMs in the clusters near the sink. Forming clusters, reducing CMs and optimal election of CH are done with respect to the position of the sink and energy level of the node in the cluster. EEEHR is efficient in providing better network lifetime and delivery of packets to the sink.

In this work, we submit energy- and memory-efficient routing protocol (E&MERP) for WSNs. A node is selected as CH based on the probability of node having high residual energy. The CH elected on high residual energy can handle various heterogeneous energy situations better than the existing clustering algorithms. When the CH is continuously used and the energy becomes equal to that of cluster members (CMs), node with higher residual energy and memory capacity will be chosen as CH. When there is a memory overflow, replacement algorithm is used and the new incoming packets are accommodated in the memory. The network lifetime will be linear with the deployed number of nodes, which guarantees better performance for increased node density in this new approach; outperforming the earlier works designed for many applications.

3 Proposed Model

E&MERP algorithm is proposed to improve the network lifetime by reducing the energy hole created because of repeatedly using the same nodes. More numbers of clusters are created near the sink associated to the numbers of clusters created far from the sink. The nodes nearer to the sink get easily burdened because these are the major access to sink. The routing of data from the same cluster and from far away cluster makes the CH nearer to the sink to get drained of its major energy. The clusters are formed in the proposed approach by considering remoteness to sink as a factor. The clusters nearer to the sink has less number of CMs because nodes nearer to the sink are burdened with its individual cluster load and routing load, whereas the clusters far away to the sink have more CMs. Equal sharing of load by the nearby clusters CH is proposed in the algorithm which extends the network lifetime. The replacement algorithm used in memory is first in first out (FIFO). As FIFO is used, priority of the message is not considered and important data may get replaced if it is there for a longer period.

Initially, the CH is nominated based on the probability of remaining energy and memory availability of the sensor node until the residual energy and available memory of CH node equals the energy of remaining nodes. Next round of selecting CH will be determined by previous election of CH in earlier rounds and the memory availability of the sensor node in WSN. Fig. 1 represents the pseudocode of this proposed protocol.

3.1 Node Deployment

In a heterogeneous network, all sensor nodes are immobile in nature and randomly deployed. All nodes have different battery levels, and the sink is driven with stable source. The base station is positioned at a fixed place inside the region of interest. The distance between the nodes can be calculated by

$$d_{xy} = \sqrt{\left(x_x - x_y\right)^2 + \left(y_x - y_y\right)^2} \tag{1}$$

In this protocol, energy consumption model is considered to be free space model to transmit in omnidirection, and it focuses on the energy used up by the nodes for communication in the network. As described in radio energy model, Eq. (2) gives the energy used up by a node to transmit the information and Eq. (4) gives the energy used up by a node to receive the information.

$$E_{sx}(k, d) = E_{diss}k + E_{free}kd^2; d < d_0 \tag{2}$$

$$= E_{diss}k + E_{multi}kd^4; d > d_0 \tag{3}$$

Input: Initial Energy of progressive node & Regular node (E_1 & E_2), Residual Energy (RE), Remoteness with sink, Memory available
Output: Best CH
Begin process:
while( )
Initial_CH Selection
CH Selection_Memory

function Initial_CH Selection
if$_1$ ***CH one hop to the sink***
calculate probable energy expense with intra cluster distance
chose node with less energy expense as Cluster Head
be a CH till node touches level E_2
else
if$_2$***(d>d*$_0$*)***
calculate probable energy expense with remoteness to sink
chose node with less energy expense as Cluster Head
end if2
e*nd if*$_1$
end function
function CH Selection_Memory
if$_3$ **RE$_1$ = E$_2$ && memory_availability>>threshold**
compute energy expenditure based on distance to sink
chose node with less energy expense as Cluster Head
end if$_3$
end function
node goes to sleep
end process( )

Fig. 1 Pseudocode for proposed routing algorithm

$$\mathrm{E_{dx}(k)} = \mathrm{E_{diss}k} \tag{4}$$

where k is the number of bits, d is distance, E_{diss} is dissipated energy for one bit to run the transceiver circuit, E_{free}, E_{multi} are dissipated energy for one bit to run the transmission amplifier based on the remoteness between the source and destination and d_0 threshold range considered by transmitter and receiver parameters.

3.2 Cluster Head Election

The energy remaining in each node is used to determine if the nodes can be used as a CM or CH. Node with higher residual energy can be used as CH, and nodes below

the range can be used as CMs. This mitigates the issue of losing a node because of energy drain and thus improves the network lifetime. CHs consume energy while transmitting, receiving and data aggregation.

Assigning the weight to the optimal probability p_{opt} of the regular nodes and the progressive nodes as p_{adv} is as follows:

Weighted probability for regular nodes and progressive nodes is given in Eqs. (5 and 6),

$$p_{\text{reg}} = \frac{p_{\text{opt}}}{1 + \alpha.m} \tag{5}$$

$$p_{\text{pro}} = \frac{p_{\text{opt}}}{1 + \alpha.m} * (1 + \alpha) \tag{6}$$

During CH election, a node already existing as CH will not become CH for one epoch, i.e. $1/p$ round if the percentage of CH equals the nodes in the cluster. The threshold for regular nodes $g\left(t_{\text{reg}}\right)$ and the threshold for progressive nodes $g\left(t_{\text{pro}}\right)$ to elect the CH in each round are given in Eqs. (7 and 8)

$$g\left(t_{\text{reg}}\right) = \begin{cases} \frac{p_{\text{reg}}}{1-\left(p_{\text{reg}} * \bmod \left(r, \text{round}\left(\frac{1}{p_{\text{reg}}}\right)\right)\right)}, & t_{\text{reg}} \epsilon G' \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

$$g\left(t_{\text{pro}}\right) = \begin{cases} \frac{p_{\text{pro}}}{1-\left(p_{\text{pro}} * \bmod \left(r, \text{round}\left(\frac{1}{p_{\text{pro}}}\right)\right)\right)}, & t_{\text{pro}} \epsilon G'' \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

where G' and G'' are the group of nodes that have not acted as CH for regular nodes and progressive nodes.

The probability of electing the nodes $\in G$ to become CHs surges in every round in the same period and becomes equal to 1 in the preceding round of the time period. The choice of selection is prepared at the start of every round by every node $\in G'$ and G'' which individually choose a random number in array of [0,1]. The node befits CH in the present round if the selected random number is less than a threshold $g\left(t_{\text{reg}}\right)$ and $g\left(t_{\text{pro}}\right)$.

Traffic load is defined as the total number of packets handled by the particular node in a particular unit time period. The number of packets generated by the node and the number of packets relayed by the node are the total number of packets handled. The load distribution for the above considered network scenario is given in [17], and the expected traffic load is expressed as

$$\text{Traffic}(r) = \frac{R^2 - r^2}{2rh}\overline{O} \tag{9}$$

where R is the network area, r is all node deployed within the network area and having distance $r \in (0, R)$, h is the mean hop length, and $\overline{O}$ is traffic rate.

Now, approximated traffic load function is given Eq. (10),

$$\text{Traffic}(r) = \begin{cases} \frac{R^2 - r^2}{2rh}\overline{O}, & \text{if } r \in (0, \text{R} - \text{h}) \\ \overline{O}, & \text{if } r \in (R - h, R) \end{cases} \tag{10}$$

The energy consumption of the network is totally depended on the total traffic load of the deployed nodes. The node density ρ in the deployed area is given as:

$$\rho = \frac{n}{\pi R^2} \tag{11}$$

The expected traffic load over the entire network nodes which are estimated by substituting Eq. (10) and (11) is given below:

$$\text{Traffic}_{\text{total}} = \iint \text{Traffic}(r)\rho d\sigma \tag{12}$$

$$\text{Traffic}_{\text{total}} \approx \frac{2R}{3h} n\overline{O} \tag{13}$$

The energy consumption of each node mainly depends on the traffic load as derived in Eq. (13). Once the node with highest residual energy is chosen, memory is considered. The replacement algorithm used in memory is first in first out (FIFO). The first packet is replaced to accommodate the packet that arrives at a router. If the packet reaches and the line (buffer space) is full, the packet gets discarded by the router which indeed reduces the throughput for a finite buffer space at the router. In order to avoid this, the packets which came first are discarded and replaced with new incoming packets. The packets are discarded without any regard that is flow of the packet or importance of the packet in the stream.

4 Simulation Results

The proposed E&MERP is compared with EEEHR, SEP, LEACH and LEACH-C algorithms which are employed with 200 nodes, and simulation bounds for the proposed algorithm are listed in Table 1. The distribution of theoretical traffic load is calculated for small network size and low density of the nodes. The experimental traffic load for the same is even and coincides with the theoretical experiments as shown in Fig. 2.

From Fig. 3, it is shown that the proposed algorithm performs better than the existing EEEHR, SEP, LEACH, LEACH-C algorithms by persisting additional number of rounds. SEP, LEACH and LEACH-C worked for 3500 rounds, whereas EEEHR lifetime exceeded 5000 rounds but had only 20 live nodes from 3000th round onward; hence, only a small network was alive for a long time.

E&MERP had more than 20 nodes till 4600 rounds, a large network was alive, and the network died at 4700th round. As E&MERP kept a larger portion of the network

Table 1 Simulation bounds

Bounds	Value
Network size	200×200 m^2
Nodes count	200
Location of base station	(100, 100)
E_e	50 nJ/bit
E_f	10pJ/bit-m^2
Initial energy	1 J
Probability of CH election	0.1
Size of data message	4000 bytes
Header size	270 bytes
Memory	2 Mb
Data rate	250 kbps

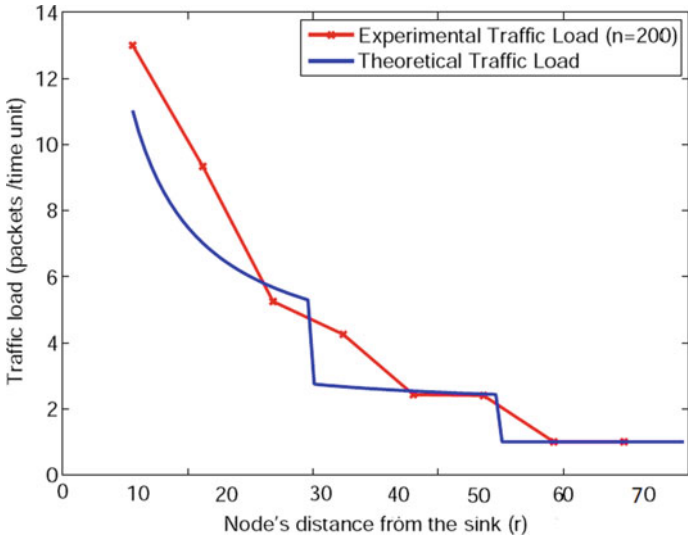


Fig. 2 Distribution of traffic load in a network

alive for a longer period relatively, this protocol is found to be better in increasing the network lifetime.

Fig. 4 shows the network lifetime of E&MERP for different values of m and alpha, where m represents the percentage of progressive nodes in the network and alpha represents the number of times the energy of progressive nodes which are greater than regular nodes. The plot indicates that as the percentage of progressive nodes and energy level of progressive node are increased, the network lifetime of E&MERP is also increased.

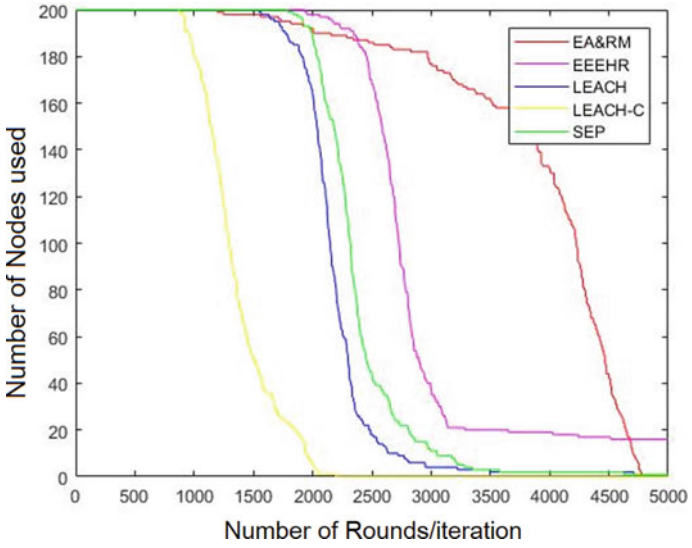


Fig. 3 Network lifetime

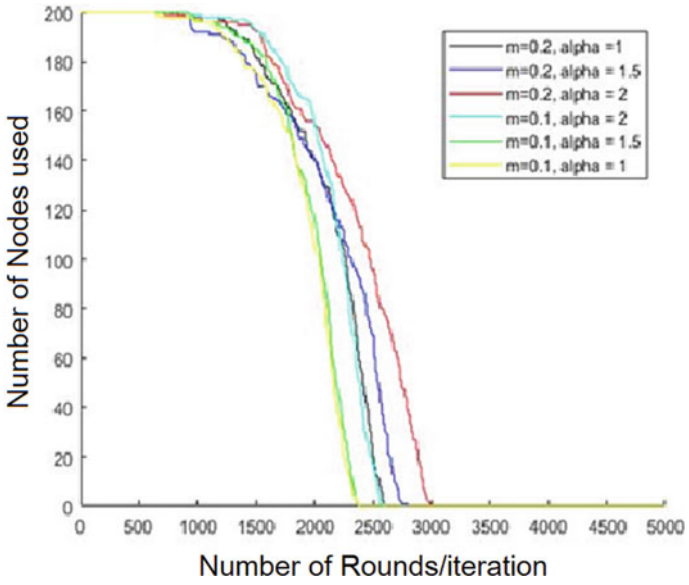


Fig. 4 E&MERP lifetime metrics for various *m* and alpha

First dead node (FDN) and half node alive (HNA) of the proposed, EEEHR, LEACH, LEACH-C and SEP algorithms are inferred from Fig. 3. The proposed algorithm FDN is at 1195 rounds, whereas the compared protocols are performing far better. The first node dies earlier before 62% of rounds. The scenario is not continuing for the HNA of the recommended algorithm. The HNA of E&MERP is 3818 rounds, improved twice the number of iterations compared to traditional LEACH and LEACH-C protocols. The HNA is 71% effective than the EEEHR and other protocols.

5 Conclusion

In this paper, E&MERP for WSN was proposed and evaluated using MATLAB. Simulation results show that E&MERP provides improved network lifetime and increased data delivery. Added load to the node makes the energy hole and hotspot problem to be critical for greater extent. The problem is solved by keeping smaller clusters near the sink and larger clusters far from the sink in the network. The performance of proposed algorithm E&MERP reduces energy hole problem and provides better clustering compared to other LEACH, LEACH-C, SEP and EEEHR algorithm. The proposed algorithm makes the network to stay alive for more than 4000 rounds and thus increasing throughput of the network. These types of network can be implemented when the sensor nodes are not permanently connected to Internet, e.g. border networks, intermittent networks. As the protocol concentrates on judicious use of energy and memory, in the long run, resources are conserved and the lifetime of network is increased substantially, thereby it serves effectively in critical situations. E&MERP helps us in a unique way for the energy problem, and buffer overflow is presently tackled by the network. Performance analysis of network for various mobility models could be done as future work.

References

1. Wu J, Dulman S, Havinga P, Nieberg T (2004) Experimental study of concurrent transmission in wireless sensor networks. In: SAFE & ProRISC-2004, pp 181–188
2. Vu TT, Nguyen VD, Nguyen HM (2014) An energy-aware routing protocol for wireless sensor networks based on k-means clustering. Recent Adv Electr Eng Relat Sci 282:297–306. https://doi.org/10.1007/978-3-642-41968-3_31
3. Jain TK, Saini DS, Bhooshan SV (2015) Lifetime optimization of a multiple sink wireless sensor network through energy balancing. J Sens 2015:1–6. https://doi.org/10.1155/2015/921250
4. Heinzelman WB, Chandrakasan AP, Balakrishnan H (2002) An application-specific protocol architecture for wireless micro sensor networks. IEEE Trans Wireless Commun 1(4):660–670. https://doi.org/10.1109/TWC.2002.804190
5. Guo P, Jiang T, Zhang K (2009) Clustering algorithm in initialization of multi-hop wireless sensor networks. IEEE Trans Wireless Commun 8(12):5713–5717. https://doi.org/10.1109/TWC.2009.12.080042

6. Kanagachidambaresan GR, Chitra A (2015) Fail safe fault tolerant mechanism for Wireless Body Sensor Network (WBSN). Wireless Pers Commun 80(1):247–260. https://doi.org/10.1007/s11277-014-2006-6
7. Mahima V, Chitra A (2017) Battery recovery based lifetime enhancement algorithm for wireless sensor network. Wirel Pers Commun 97(4):6541–6557. https://doi.org/10.1007/s11277-017-4854-3
8. Lindsey S, Raghavendra C (2002) PEGASIS: power-efficient gathering in sensor information systems. In: Proceedings of the IEEE aerospace conference, pp. 1125–1130. https://doi.org/10.1109/AERO.2002.1035242
9. Smaragdakis G, Matta I, Bestavros A (2004) SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor networks. In: Second international workshop on sensor and actuator network protocols and applications (SANPA 2004). hdl.handle.net/2144/1548
10. Soro S, Heinzelman W (2005) Prolonging the lifetime of wireless sensor networks via unequal clustering. In: Proceedings of the 19th IEEE international parallel and distributed processing symposium (IPDPS), pp 236–240. https://doi.org/10.1109/IPDPS.2005.365
11. Saranya V, Shankar S, Kanagachidambaresan GR (2018) Energy efficient clustering scheme for wireless sensor network with mobile sink. Wireless Pers Commun 100(4):1553–1567. https://doi.org/10.1007/s11277-018-5653-1
12. Azharuddin M, Jana PK (2015) A distributed algorithm for energy efficient and fault tolerant routing in wireless sensor networks. Wireless Netw 21(1):251–267. https://doi.org/10.1007/s11276-014-0782-2
13. SarmaDhulipala VR, Kanagachidambaresan GR, Chandrasekaran RM (2012) Lack of power avoidance: a fault classification based fault tolerant framework solution for lifetime enhancement and reliable communication in wireless sensor network. Inf Technol J 11(6):719–724. https://doi.org/10.3923/itj.2012.719.724
14. Xiaohan L, Xiaoyu J, Xinyan Z, Longdao C (2018) Efficient link-delay aware routing in wireless sensor networks. IEEE Sens J 18(2):837–848. https://doi.org/10.1109/JSEN.2017.2772321
15. Korhan C, Tamer D (2018) Energy aware multi-hop routing protocol for WSNs. IEEE Access J 6:2622–2633. https://doi.org/10.1109/ACCESS.2017.2784542
16. Prabha KL, Selvan S (2018) Energy efficient energy hole repelling (EEEHR) algorithm for delay tolerant wireless sensor network. Wireless Pers Commun 101(3):1395–1409. https://doi.org/10.1007/s11277-018-5768-4
17. Wang Q (2010) Traffic analysis & modeling in wireless sensor networks and their applications on network optimization and anomaly detection. Netw Protoc Algorithms, Macrothing Inst 2(1):74–92. https://doi.org/10.5296/npa.v2i1.328

Multi-scale Aerial Object Detection Using Feature Pyramid Networks

Dennis George Johnson, Nandan Bhat, K. R. Akshatha, A. K. Karunakar, and B. Satish Shenoy

Abstract Aerial object detection on an UAV or embedded vision platform requires accurate detection of objects with various spatial scales and has numerous applications in surveillance, traffic monitoring, search, and rescue, etc. The task of small-object detection becomes harder while using standard convolutional neural network architectures due to the reduction in spatial resolution. This work evaluates the effectiveness of using feature pyramid hierarchies with the Faster R-CNN algorithm for aerial object detection. The VisDrone aerial object detection dataset with ten object classes has been utilized to develop a Faster R-CNN ResNet model with C4 and FPN architectures to compare the performance. Significant improvement in the performance obtained by using feature pyramid networks for all object categories highlights their importance in the multi-scale aerial object detection task.

Keywords Object detection · Feature pyramid networks · Aerial images · VisDrone · Faster R-CNN

D. G. Johnson · N. Bhat
Department of Computer Science Engineering, Manipal, India

K. R. Akshatha (✉)
Department of Electronics and Communication Engineering, Centre of Excellence in Avionics and Navigation, Manipal, India
e-mail: akshatha.kr@manipal.edu

A. K. Karunakar
Department of Data Science and Computer Applications, Manipal, India

B. Satish Shenoy
Department of Aeronautical and Automobile Engineering, Centre of Excellence in Avionics and Navigation, Manipal, India

D. G. Johnson · N. Bhat · K. R. Akshatha · A. K. Karunakar · B. Satish Shenoy
Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_31

1 Introduction

Object detection is the main task in computer vision that involves detection of object instances in digital images or videos. It aims at finding the position of objects within the image (object localization) and determining the class to which they belong (object classification). Deep learning has been the biggest factor in the progress of object detection while working in conjunction with traditional computer vision and image processing techniques or by supplanting them completely [14]. Substantial work has been done on generic object detection, which is broadly classified into two categories. The first one follows a traditional two-step approach where the region proposal is generated in the first step, which is followed by a classification of the proposed regions in the next step. The second approach makes use of a unified framework where both regression of the bounding box and classification are achieved directly without having a separate region proposal step. Some of the two-stage approaches are R-CNN [6], SPP-net [8], Fast R-CNN [5], Faster R-CNN [21], R-FCN [3], etc. Some of the regression/classification-based frameworks are YOLO [2, 18–20], Single Shot Multi-Box Detector (SSD) [15], Deconvolutional Single Shot Multi-Box Detector (DSSD) [4], etc.

In recent years, advancement in technologies and the cost reduction of Unmanned Aerial Vehicles (UAVs) facilitated their use for civil applications and not just military uses. Among the pool of UAVs, low altitude UAVs like drones are preferable for computer vision applications as they can get clear and non-blurry images. The UAVs have high mobility and can cover a wide field of view and therefore can be effectively used in gathering data. Some of its applications include agricultural production, environmental monitoring, fast delivery, and aerial photography, and many more. The aerial photography-related applications include surveillance, geophysical exploration, 3D reconstruction, disaster relief and rescue, mapping, etc. [24]. Illumination variations, cluttered backgrounds, and varying pose and size of the objects are major challenges that arise in UAV-based object detection. UAV footage often has objects of various scales, small objects being one of them. Many of the algorithms developed for generic object detection perform well on detecting medium to large objects, but struggle to detect objects of small or varying scales. In this work, we evaluate the importance of using Feature Pyramid Networks (FPN) with the presence of a region proposal network for detecting multi-scale objects present in aerial scenes [12]. Multi-scale hierarchies help to associate rich semantic features with spatial information at all levels.

2 Related Work

Aerial object detection is a challenging task of computer vision being explored by various techniques that are successful in general object detection. After the initial use of CNNs in aerial object recognition, Sevo and Avramovic introduced CNN into aerial object detection and it was shown to be effective [23]. To preserve the

low-level features for small-object detection, Tayara and Chong used a single-stage densely connected feature pyramid network for high-resolution aerial images [29]. Feature fusion deep network (FFDN) method preserves not only the hierarchical feature representation for small target detection but retains the spatial relationship between the objects within the images [17]. Multi-branch Parallel FPN (MPFPN) attempts to extract more information of small objects from aerial images by using parallel branches to recover features from shallow layers [16]. Xiao et al. proposed combining deep residual network, FPN, and rotation region proposal networks to extract multi-scale features and generate the aligned bounding boxes around the candidate regions [32]. SPB-YOLO is another method that deals with countering difficulties because of the scale variations and dense distribution of various objects present in aerial scenes [30]. The adaptive feature selection module introduces a fusion approach for multi-level representations enhancing important clues in certain feature levels in combination with the class-aware sampling mechanism [7]. Sun et al. introduced a novel loss function called salience biased loss, where the input training examples are treated differently based on the complexity of the images [27]. Adaptive Saliency biased loss being another method of extracting salient information from aerial images which can be applied at either the image level or the anchor level as in [28]. Anchor optimization methods as suggested by Ahmad et al. aims to deal with the intrinsic class imbalance problem in aerial images where background samples dominate object samples purely by using anchor optimization [1]. Hou et al. proposed an object detection framework that uses sample balance strategies and weight GANs sub-networks for detection from drone imagery [10]. In urban environments, where vehicles are one of the most prevalent object classes, methods like orientation-aware vehicle detection in aerial images include other techniques like feature fusion [25]. Furthermore, Multi-scale Keypoint Detection Network (MKD-Net), after fusing the features, generates two feature maps at different scales to effectively detect object instances of different scale variations in aerial images [26]. Sambolek and Ivasic-Kos have shown the performance of various model architectures for automatic person detection in aerial images for search and rescue process [22]. Other niche works like trash and litter detection from aerial images might be taken into consideration [11].

3 Methods

The adopted methodology has been outlined in this section, which includes dataset collection, preprocessing of data, training, and evaluation of the object detection models. The standard online available VisDrone object detection dataset [33] is collected. The correctness of the given annotations for ten object categories, i.e. pedestrian, people, bicycle, car, van, truck, tricycle, awning tricycle, bus and motor given in text format, are checked before training the model. The train-set images are used to build the Faster R-CNN model using similar backbones with C4 and FPN architectures. The performance evaluation of the trained models has been performed on the test-dev dataset using standard COCO evaluation metrics [13].

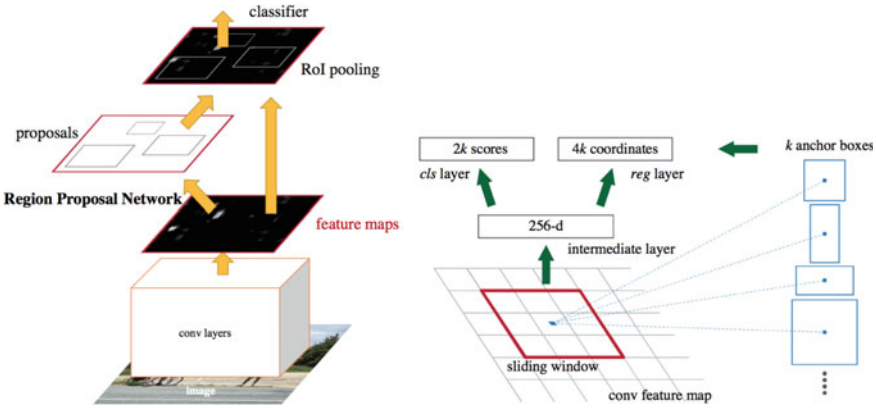


Fig. 1 Faster R-CNN approach for object detection [21]

3.1 Faster R-CNN

A Faster R-CNN algorithm is used in this research work due to its good accuracy in object detection. The Faster R-CNN architecture uses two modules. The first one is the region proposal network (RPN), that uses a deep fully convolutional backbone network for region proposal generation. The second module is the Fast R-CNN detector, that utilizes the generated regions by RPN to detect the presence of the objects. The region proposal network works on an image of arbitrary size and gives out the objectness score for each object proposals to estimate the presence of an object in the region. The feature maps from the same convolutional backbone structure are used by both RPN and the Fast R-CNN detector. The RPN slides a small network over the feature maps from the shared convolution backbone and then parses the output through two fully connected layers, i.e. box-regression layer and a box-classification layer. The Faster R-CNN approach for object detection is shown in Fig. 1.

Additionally, anchor boxes are used at each location in the RPN. The anchors are translation invariant and are useful in addressing the issue of varying scales and aspect ratios. This enables the classification and regression of bounding boxes concerning anchor boxes of multiple scales and aspect rations while utilizing input (image and feature map) of a specific size. The features for RPN and the Fast R-CNN detector are shared in Faster R-CNN.

3.2 Feature Pyramid Networks [12]

Objects in aerial images generally have variation in scale and detecting them becomes harder while using standard convolutional neural network architectures due to the reduction in spatial resolution. The use of convolutional neural networks with multi-scale hierarchy helps to combine rich semantic information with spatial

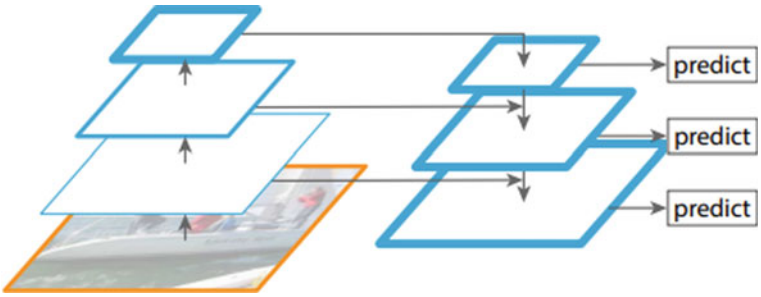


Fig. 2 Feature pyramid networks in object detection [12]

information at all levels. Thus feature pyramid networks are used in this work for performing aerial object detection tasks. The ConvNet's pyramidal feature hierarchy is utilized efficiently in the Feature Pyramid Network. The architecture of the FPN is shown in Fig. 2.

The construction of the pyramid involves bottom-up, top-down, and lateral connections. The FPN structure is created by combining one feature map from each stage (the continuous layers that produce the same dimensional feature maps belong to the same stage) to form the backbone structure. The last feature map from each stage is used in the generation of the pyramid. The top-down pathway and the bottom-up pathway features are enhanced with lateral connections.

We use Faster R-CNN model architecture that utilizes similar backbones to compare the performance with and without FPN. Faster-RCNN utilizes a 101-layer residual network (ResNet-101) and a 50-layer residual network (ResNet-50) as backbones in the study. These backbone networks are utilized to extract feature maps from the input images at different scales. Using ResNets as backbones allows for training deeper networks while avoiding the performance degradation problem associated with them [9]. This is achieved by using shortcut connections that allow for skipping one or more layers. This feature of a ResNet helps mitigate the vanishing/exploding gradient problem. The shortcut connections provide an identity mapping and thus facilitate easier optimization and better accuracy for deeper networks. Two backbone configurations of ResNets are used to compare the effectiveness of hierarchical feature pyramids.

3.3 ResNet with C4

This uses a backbone that ends at the conv4 layer and a box prediction head that consists of the conv5 layer of a ResNet. This configuration follows the original ResNet paper implementation for both ResNet-50 and ResNet-101 backbones [9]. These backbones differ by the number of residual blocks that are employed within their conv4 layers.

3.4 ResNet with FPN

ResNet FPN backbone with standard convolutional and fully connected heads are used for bounding box prediction. This configuration uses the layers conv2–conv5 of the ResNet for the bottom-up pathway of FPN. Each lateral connection merges the feature maps of similar spatial scales from the bottom-up pathway and top-down pathway. These pathways are merged by element-wise addition at each scale.

3.5 Evaluation Metrics

Average precision is used to quantify the accuracy of the model trained. Precision measures the percentage of the predictions that are correct and recall measures how well the positives are found. Average precision is a metric that combines precision and recall, it is the mean of the precision scores for all classes. Additionally, intersection over union is used to measure overlap between bounding boxes. A threshold is set to determine if the predicted bounding box is a true positive or a false positive. COCO evaluation metrics have been used in this work to evaluate the trained models [13].

4 Experiments and Results

The experiment has been carried out using a single Nvidia Tesla K80 GPU provided by Google Colaboratory. The object detection models are developed with the help of the PyTorch-based Detectron2 library [31].

4.1 Dataset Details

The VisDrone-DET 2021 dataset has been used to train and evaluate the model. The images in this dataset are acquired from drone-mounted cameras, in 14 different cities of China, having a variety of objects, and showing the varied density of objects. The dataset consists of 10,209 images taken in unconstrained complex scenarios. The dataset comprises of 6,471 training images, 548 validation images and 1610 images in test-dev set [33]. The dataset is fully annotated for 10 classes: pedestrian, people, bicycle, car, van, truck, tricycle, awning tricycle, bus and motor in text format.

4.2 Experimental Results

Transfer learning is employed and the pre-trained models of the MS COCO dataset are used for training. The Faster R-CNN model performance using similar back-

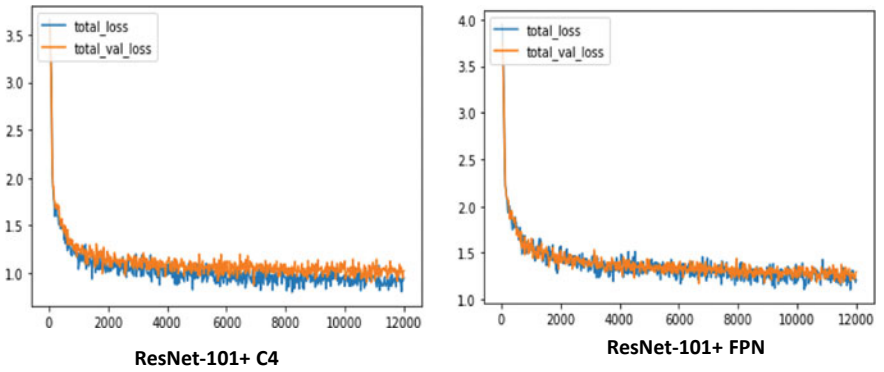


Fig. 3 Training and validation loss of ResNet-101 with a C4 backbone (left) and ResNet-101 with FPN backbone (right) in object detection

bones is compared with and without FPN. We evaluate using Faster R-CNN that utilizes a 101-layer residual network (ResNet-101) and a 50-layer residual network (ResNet-50) as backbones. Fine-tuning is carried out through all layers using the VisDrone-DET 2021 training dataset and a base learning rate of 2.5e−4 and a batch size of 2 on a single GPU. Each model was trained for 12,000 iterations while using the VisDrone-DET validation dataset to monitor convergence and avoid overfitting. Figure 3 shows training loss (blue) plotted against the validation loss (orange) measured during training for ResNet-101 with a C4 backbone (left) and ResNet-101 with FPN backbone (right).

A minimum score threshold of 0.6 was chosen to achieve high recall with less false positives and to reduce the inference time that involves post-processing steps like non-maximum suppression. Two backbone configurations for ResNets, i.e. one with C4 and the other with FPN are used to compare the effectiveness of hierarchical feature pyramids. Table 1 shows the evaluation results of various models on the VisDrone dataset and their average detection time.

From Table 1 we can infer that, FPN out-performed the C4 architecture for both ResNet-50 and ResNet-101 backbones in terms of all metrices of COCO evaluation scheme. The class-wise performance obtained for these models is shown in Table 2.

Table 1 Evaluation results on VisDrone-DET 2021 test-dev dataset

Backbone	AP	AP50	AP75	APs	APm	APl	Inference time (s)
R101 + FPN	14.19	26.48	13.72	5.64	22.74	36.42	0.161
R101 + C4	12.25	22.95	11.75	3.98	19.83	36.42	0.500
R50 + FPN	13.16	25.19	12.61	5.47	20.9	32.10	0.117
R50 + C4	11.62	21.31	11.48	3.88	18.77	32.17	0.450

Table 2 Class-wise AP results on test-dev dataset

Backbone	Bicycle	Truck	Bus	Pedestrian	Car	Tricycle	Motor	People	Van	Awning tricycle
R101 + FPN	3.41	18.26	34.03	10.76	37.11	5.56	6.32	3.82	18.29	4.34
R101 + C4	2.23	14.35	33.81	7.34	33.02	4.32	5.53	2.44	15.80	3.64
R50 + FPN	3.03	14.48	31.27	10.38	36.09	4.74	6.13	3.39	18.18	3.88
R50 + C4	2.42	13.89	30.46	7.77	35.10	3.45	4.52	1.99	14.57	2.07

Fig. 4 Detection results on VisDrone test-dev dataset. C4 configuration (left) and FPN (right)

The AP for object categories, i.e. the car, bus, van, etc. is quite high because of their large size in the images and the availability of more object instances for these categories. Though FPN improved the performance for all object categories, there is a scope for further improvement. Figure 4 shows the sample predictions obtained for a random image from the test-dev dataset.

4.3 Discussion

The feature pyramid networks have shown significant improvement in the performance for Faster R-CNN with both ResNet-50 and ResNet-101 backbones when compared to the C4 network without using FPN. The improvement in the performance is seen for all the object categories present in the VisDrone dataset. The feature pyramid networks utilized lesser inference time, which is another advantage for using these networks. By adapting the Region Proposal Network with FPN, it is not necessary to have multi-scale anchor boxes for each feature map. This differs from the conv5 head used by the C4 backbone configuration which proposes multi-scale anchor boxes from the single final feature map. We can observe faster inference times while using FPN for region proposals as a single scale anchor box is assigned to each level of the hierarchy. For each level, multiple aspect ratios are used for proposals 1:1, 1:2, 2:1. Features for both the detection and region proposal network are shared, which further reduces inference time.

5 Conclusion

We have demonstrated the effectiveness of using feature pyramid networks for the task of aerial object detection. The use of a feature pyramid network reveals significant improvements over the baseline model configurations in overall detection accuracy due to the ability to use rich semantics at all levels with marginal extra cost. This performance improvement emphasizes the viability of such architectures

for multi-scale object detection on a UAV platform. This warrants future study into developing more efficient deep-learning architectures for improving the performance of small-scale object detection on UAVs and other embedded vision platforms.

References

1. Ahmad M, Abdullah M, Han D (2020) Small object detection in aerial imagery using RetinaNet with anchor optimization. In: 2020 international conference on electronics, information, and communication (ICEIC). IEEE, pp 1–3
2. Bochkovskiy A, Wang CY, Liao HYM (2020) Yolov4: optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934
3. Dai J, Li Y, He K, Sun J (2016) R-FCN: object detection via region-based fully convolutional networks. In: Advances in neural information processing systems, pp 379–387
4. Fu CY, Liu W, Ranga A, Tyagi A, Berg AC (2017) DSSD: Deconvolutional single shot detector. arXiv preprint arXiv:1701.06659
5. Girshick R (2015) Fast R-CNN. In: Proceedings of the IEEE international conference on computer vision (ICCV)
6. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 580–587
7. Gong Z, Li D (2020) Towards better object detection in scale variation with adaptive feature selection. arXiv preprint arXiv:2012.03265
8. He K, Zhang X, Ren S, Sun J (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans. Pattern Anal. Mach. Intell. 37(9):1904–1916
9. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
10. Hou X, Zhang K, Xu J, Huang W, Yu X, Xu H (2021) Object detection in drone imagery via sample balance strategies and local feature enhancement. Appl. Sci. 11(8):3547
11. Kraft M, Piechocki M, Ptak B, Walas K (2021) Autonomous, onboard vision-based trash and litter detection in low altitude aerial images collected by an unmanned aerial vehicle. Remote Sens. 13(5):965
12. Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S (2017) Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2117–2125
13. Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL (2014) Microsoft COCO: common objects in context. In: European conference on computer vision. Springer, pp 740–755
14. Liu L, Ouyang W, Wang X, Fieguth P, Chen J, Liu X, Pietikäinen M (2020) Deep learning for generic object detection: a survey. Int. J. Comput. Vis. 128(2):261–318
15. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC (2016) SSD: single shot multibox detector. In: European conference on computer vision. Springer, pp 21–37
16. Liu Y, Yang F, Hu P (2020) Small-object detection in UAV-captured images via multi-branch parallel feature pyramid networks. IEEE Access 8:145740–145750
17. Long H, Chung Y, Liu Z, Bu S (2019) Object detection in aerial images using feature fusion deep networks. IEEE Access 7:30980–30990
18. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788
19. Redmon J, Farhadi A (2017) Yolo9000: better, faster, stronger. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7263–7271

20. Redmon J, Farhadi A (2018) Yolov3: an incremental improvement. arXiv preprint arXiv:1804.02767
21. Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: towards real-time object detection with region proposal networks. In: Advances in neural information processing systems, pp 91–99
22. Sambolek S, Ivasic-Kos M (2021) Automatic person detection in search and rescue operations using deep CNN detectors. IEEE Access 9:37905–37922
23. Ševo I, Avramović A (2016) Convolutional neural network based automatic object detection on aerial images. IEEE Geosci Remote Sens Lett 13(5):740–744
24. Shakhatreh H, Sawalmeh AH, Al-Fuqaha A, Dou Z, Almaita E, Khalil I, Othman NS, Khreishah A, Guizani M (2019) Unmanned aerial vehicles (UAVs): a survey on civil applications and key research challenges. IEEE Access 7:48572–48634
25. Shi F, Zhang T, Zhang T (2020) Orientation-aware vehicle detection in aerial images via an anchor-free object detection approach. IEEE Trans Geosci Remote Sens 59(6):5221–5233
26. Su J, Liao J, Gu D, Wang Z, Cai G (2020) Object detection in aerial images using a multiscale keypoint detection network. IEEE J Sel Top Appl Earth Obs Remote Sens 14:1389–1398
27. Sun P, Chen G, Luke G, Shang Y (2018) Salience biased loss for object detection in aerial images. arXiv preprint arXiv:1810.08103
28. Sun P, Chen G, Shang Y (2020) Adaptive saliency biased loss for object detection in aerial images. IEEE Trans Geosci Remote Sens 58(10):7154–7165
29. Tayara H, Chong KT (2018) Object detection in very high-resolution aerial images using one-stage densely connected feature pyramid network. Sensors 18(10):3341
30. Wang X, Li W, Guo W, Cao K (2021) SPB-YOLO: an efficient real-time detector for unmanned aerial vehicle images. In: 2021 international conference on artificial intelligence in information and communication (ICAIIC). IEEE, pp 99–104
31. Wu Y, Kirillov A, Massa F, Lo WY, Girshick R (2019) Detectron2. https://github.com/facebookresearch/detectron2
32. Xiao J, Zhang S, Dai Y, Jiang Z, Yi B, Xu C (2020) Multiclass object detection in UAV images based on rotation region network. IEEE J Miniat Air Space Syst 1(3):188–196
33. Zhu P, Wen L, Du D, Bian X, Hu Q, Ling H (2020) Vision meets drones: past, present and future. arXiv preprint arXiv:2001.06303

Application-Based Approach to Suggest the Alternatives of Daily Used Products

Vedant Dokania, Pratham Barve, Krushil Mehta, Dev Patel, and Mayur M. Sevak

Abstract For decades, we consider nature conservation a serious topic. However, with the support of people of all ages, one can excitingly hold nature. A multipurpose application is created using flutter and firebase machine learning vision (ML vision) in which one can use a camera/gallery image to find the alternatives of the day-to-day required items and can also see the amount of pollution caused by a similar item. The application also suggests an environment-friendly alternative to the daily used products. Soon, detecting items and seeking alternatives for the same will become a fun learning task. One can also share the amount of pollution they have saved, after selecting the environment-friendly alternatives and can also win rewards and social media badges. The application also has a community tab in which one can share their thoughts regarding nature conservation and a group of people can conduct a healthy discussion on the nature conservation topics.

Keywords Machine learning · Firebase · Flutter · Environment-friendly · Pollutant

1 Introduction

The aim is to achieve a greener world where all the creatures live in utter peace and harmony, a better place for our future generations to stay. A balanced ecosystem is very important for the sustenance of the world. An ecosystem where its members are dependent on each other but none of them harms the other counterparts just for their benefit and luxury. The mobile application is developed that suggests the alternatives of the daily used harmful products which leave a hazardous impact on nature and the environment. The application enables the user to capture a picture or use the existing ones from the mobile's gallery. The objects will then be detected and recognized using Google ML Vision Toolkit. The alternative of the same will be suggested along with the pollution caused by both—the object detected and the alternative, which makes it easy for the user to compare them. It also gives us the

V. Dokania · P. Barve · K. Mehta · D. Patel · M. M. Sevak (✉)
Birla Vishvakarma Mahavidyalaya Engineering College, Anand, Gujarat, India
e-mail: mayur.sevak@bvmengineering.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_32

feature of tracking the location of the mobile phone and informing about the pollution causing elements at the same location. One can also write notes in the To-Do tab. This tab will also have the tasks that the user is supposed to do by adopting the greener replacements. The user needs to register and securely log in to the application before availing of the above-mentioned benefits.

2 Technology Used

2.1 Firebase ML Vision

Firebase ML kit [1] is developed by Google which supports various modules like text recognition, face recognition, barcode scanning, image scanning, and many more. The application designed here makes use of an Image labeler. Image labeler can detect the object based on the training dataset provided. Firebase ML vision provides two different options to detect the object: on-device and cloud. On on-device one needs to download the firebase ML vision SDK to the device and device computational power is being used to detect the object. In on cloud mode, the firebase cloud platform only takes care of detecting the image. Firebase already contains a dataset of most of the objects available and Google updates the same on regular basis. Along with the training dataset provided by Google one can also insert the objects with their labels and upload the file on the cloud portal and then firebase trains the data on its own using the user dataset.

2.2 Flutter—A Cross-Platform Framework

Flutter [2] is a cross-platform application development framework developed by Google. On a single code-based, we can make the android application, iOS application, desktop application, and web application. The flutter plugin is installed on the android studio. Flutter supports various libraries created by the user and also developed by the flutter team. The libraries used in the application are flutter_firebase, flutter_firestore, flutter_text to speech. This library helps the developer to connect the application with the backend that is Google firebase.

3 Pollution Causing Objects

The fact that pollution continues to rise has become a well-known reality. The term "pollution" refers to the presence of any unwanted foreign substance in a system. The pollution is caused because of numerous pollutants which result in contamination

of natural resources. This is mostly due to human activities, which degrade the ecosystem in a variety of ways. Some pollution creating objects are:

- Car

Vehicle emissions have an impact on both local and regional air quality. The vehicle emits various gases like nitrogen oxides, hydrocarbons, carbon monoxide, and many more. To avoid the same one needs to use the electric vehicle in day-to-day life.

- Plastic

Plastic pollution has risen to the top of the environmental priority list. From around two and a half million tons in 1950 to around 450 million tons in 2015, production grew increases at an exponential rate [3]. By 2050, production is predicted to double. Most of the plastic ever made is in last 15 years, also as per the data approximately 7–8 million tons of plastic garbage dump by developing countries each year in oceans. Garbage bags full of plastic trash which is equivalent to one football ground, dump it into the oceans every year. Plastics have some qualities like flexibleness, long lasting and strong enough to carry weight, to achieve these characteristics extra additives and chemicals are used. Thus, plastic cannot degrade or decompose before 300–400 years.

- Air Conditioner and Refrigerator

HFCs, or hydrofluorocarbons, are chemical compounds that keep the temperature cool in air conditioners and freezers on hot summer days. HFCs contribute to global warming in the same way that other greenhouse gases do. HFCs account for a small percentage of total greenhouse gas emissions, but they are thousands of times hazardous than carbon dioxide in terms of trapping the heat in the atmosphere.

- Aircraft

Nitrogen dioxide (NO_2) is the primary pollutant of concern for airports. Nitrogen oxide (NO_x) is emitted from aircraft, surface traffic, and airport operations. NO_x in the lower atmosphere hinders the process of making ozone thus results in global warming.

3.1 NASA Satellite Data

The NASA satellite [4] provides the data of population in a geographic region constantly. Below is the list of some satellites which provides the data regarding pollution levels:

- AURA

The Aura has four research instruments for the measurement of ozone trends, changes in the air quality which are alimentally responsible for climate change.

- TERRA

The Terra has the label of first earth observation system, it is collecting data of the land, atmosphere, oceans, and solar radiation.

- LANDSAT

The Landsat has two satellites Landsat 5 and Landsat 7 that provide information for remote sensing, ground system development and Geographic Information System (GIS) science and applications [5].

3.2 List of Objects with Their Safer Alternatives

Branded products are often thought of as being safe and hygienic although it is not the case. Several goods which we utilize in our daily chores comprises various hazardous substances which may damage our environment if they are not disposed of properly.

If people can realize and distinguish between hazardous and non-hazardous products, can specify their alternatives and help save the environment [6]. Various ecolabels are implanted on the packet of any product which provides information about whether it meets specific environmental criteria to the consumers. For instance, a label that reads "100% recycled" tells a person that a product has been made from recycled materials. These labels provide a calculation of the CO_2 production derived from the production of the product till it is disposed of. Flammable, explosive, toxic, dangerous, or warning are some adjectives used on these labels to seek the consumer's attention on it [7]. Instead of merely looking at the label user can use the application which will detect the product and indeed will categorize the same on different parameters and will also suggest the alternative for the same. Table 1 is the data for many such replacements that can be deployed.

4 System Workflow

Firstly the user needs to register himself in the application. Then user encounters various sections like community tab, upload/capture, to/do tab, and so on. Based on the preference user can select any of the tabs. After selecting the upload/capture tab user needs to take picture of the image from the camera application or can upload the same from the gallery. After successfully uploading the image, firebase ML vision will detect the object and label the same. After searching about the label from the alternative object array, the application will suggest the proper alternatives and will also give information about the amount of pollution created by the object. Users can than share the same information in the community tab and also discuss with other users regarding the alternatives (Figs. 1, 2 and 3).

Table 1 List of safer alternatives

Substance	Potential hazard	Safer alternative
Plastic glass	Flammable, toxic	Metal glass
Plastic toys	Toxic	Soft toys/wooden toys
Plastic ball pens	Toxic	Pencils
Polythene bags	Toxic	Cloth/paper bags
Syringe	Toxic	Needle-free injectors
Razor blades		Electric trimmers scissors
Air conditioner	CFC/HFC	Conventional fans
Motorbike/cars	Emits CO_2	Electric/solar-based vehicles or conventional bicycles
Laundry detergents	Phosphates, surfactants in them	Choose phosphate-free options
Toothpaste/soaps/facewash	Contains micro leads	Check quality before buying
Tea bags	Polypropylene	Sustainable tea bags
Plastic bottles	Non-biodegradable	Metal/glass bottles

4.1 Some Features of the Application

1. The application is designed using a cross-platform framework, i.e., flutter. The application can work on android devices, iOS devices and also can be used as a web application.
2. The application supports secure login with different social media handles like Google, Facebook, Twitter and many more.
3. The application is designed in such a way that it can easily be used by the persons of all age groups and is indeed a user friendly.
4. The performance of the application is increased as Google firebase takes care of the machine learning part and would suggest the alternatives based on the image detected.
5. The simplicity and interaction with community helps to scale the application to large audience.

5 Conclusion

Environmental pollution is not a replacement development, nevertheless, it remains the world's greatest drawback facing humanity and also the leading environmental cause of morbidity and mortality. We need to understand the fragility of our environment and the importance of its protection. We need to protect the environment for our future generations to give them a chance to take benefits from nature and use the same for their livelihood. That's why an idea is put forward about developing a mobile application that suggests the alternatives of the daily used harmful products

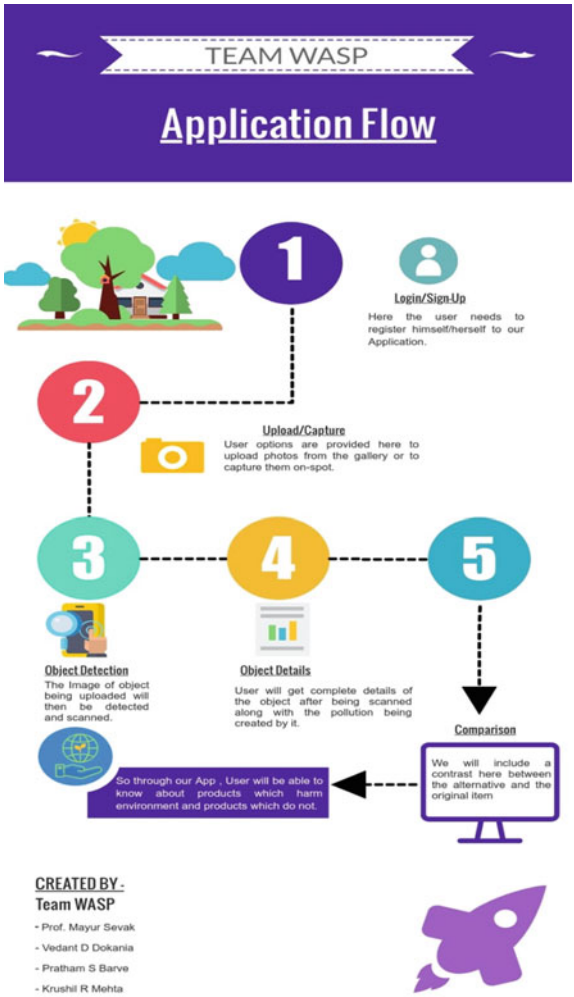


Fig. 1 Process workflow of the application

that leave a hazardous impact on nature and the environment. Here the user has to take a snapshot of any object that one wants information about. One can also upload the image from their device's gallery. The application then scans the image and then provides necessary statistics and information about it. In this way, the users can also know their contribution toward reducing pollution.

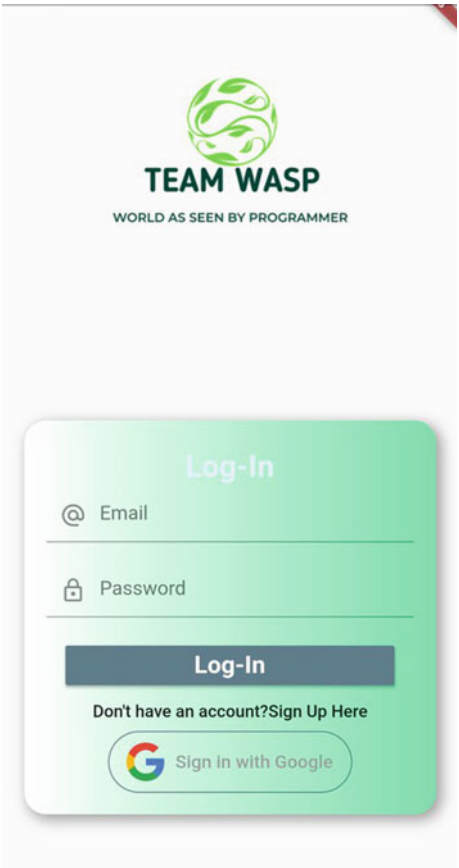


Fig. 2 Application login page

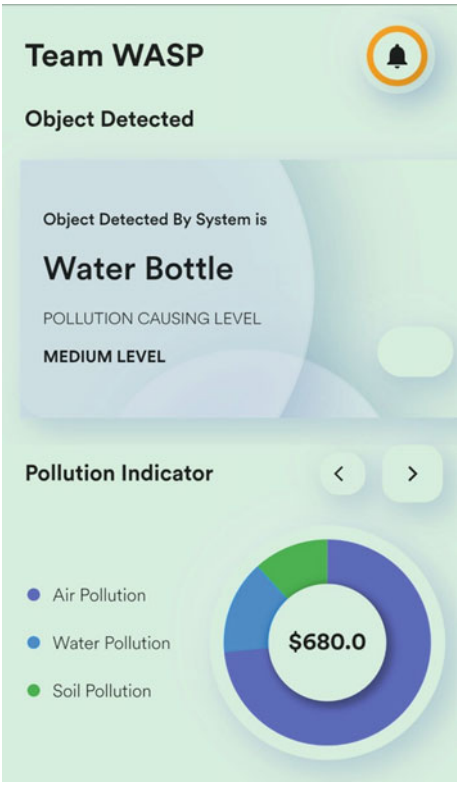


Fig. 3 Application page showcasing the pollution level caused by the plastic water bottle

Acknowledgements We would endorse the Electronics and Communication department of Birla Vishvakarma Mahavidyalaya College for allowing us to exhibit our skills. Furthermore, we are extremely grateful to NASA International Space Apps Challenge (an International Level Hackathon) [8] held from 2nd October 2020–4th October 2020, for selecting us as a finalist team for Gujarat Region and allowing us to manifest our ideas.

References

1. Firebase ML kit. https://firebase.google.com/docs/ml-kit
2. Flutter by Google. https://flutter.dev/
3. The world's plastic pollution crisis explained by national geographic. https://www.nationalgeographic.com/environment/article/plastic-pollution. Accessed 28 Aug 2021
4. NASA climate change homepage. https://climate.nasa.gov/. Accessed 30 Aug 2021
5. NASA visible earth webpage. https://eospso.nasa.gov/content/all-missions. Accessed 30 Aug 2021

6. Dietz T, Shwom RL, Whitley CT (2020) Annual climate, change and society. Rev Sociol 46(1):135–158
7. Calzadilla A, Rehdanz K, Betts R et al (2013) Climate change impacts on global agriculture. Clim Change 120:357–374
8. NASA space apps challenge website. https://2020.spaceappschallenge.org/

Comparative Analysis of Multi-level Inverters with Various PWM Techniques

V. Ramu, P. Satish Kumar, and G. N. Srinivas

Abstract Multi-level inverter (MLI) with multi voltage level has consistently been common in inverter arrangements because of its added benefits of low distortion with improved output wave, low-esteem filter segments, higher power yield, etc. This paper hits on with the comparative analysis of 17-level voltage source inverter with PWM techniques to the conventional topologies of cascaded H Bridge MLI. The two inverters are contrasted in terms of THD, output voltage and currents. The switching instances ensure the minimum number of switching transition in contrast to the H Bridge MLI and allows us to design with reduced switch count. The proposed 17 level MLI is verified by required simulations in PWM techniques and cascaded H bridge model on MATLAB and the corresponding sim-results are presented.

Keywords 17 level inverter · Cascaded H bridge (CHB) MLI · PWM techniques · Switching instances

1 Introduction

The rapid development in the modern power electronics has paved a multi-level inverters as newfangled conversion approach. The interconnection of switches with low count attaining high power output and efficiency with higher quality of power and less distortion in multi-levels has drawn attention in power conditioning devices.

V. Ramu (✉) · G. N. Srinivas
Jawaharlal Nehru Technological University Hyderabad, Hyderabad, India
e-mail: ramu.765@kakatiya.ac.in

G. N. Srinivas
e-mail: gnsgns.srinivas785@jntuh.ac.in

P. Satish Kumar
University College of Engineering, Osmania University, Hyderabad, India
e-mail: satish_8020@osmania.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_33

To achieve high voltage and high power within switches and to work on output waves, staggered Voltage source inverters (VSIs), in light of diode-clipped, flying-capacitor, cascaded bridges which are proposed for supplanting the low-level converters with a series association of switches and also transformer coupled inverters in medium voltage and high voltage (HV) applications, for example, engine drives' and static compensators [1–5]. Multi-level inverters advancements have alluring provisions such as high voltage capability, decreased normal mode voltages are close to sinusoidal and more modest or even zero output channels making them appropriate for use in high power applications and distributed generators.

As of the conventional topologies in multi-level inverters we are provided with neutral clamped, capacitor clamped, cascaded H Bridge which comes with drawbacks of unequal voltage sharings that give arise to increase in number of clamping diodes. Fly bridge capacitor model though have several attractive properties such as transformerless operation and reduced switching stress among the power switching devices but limited by usage of capacitors for high voltage levels.

In continuation to this a double FS MLI in [6] was designed by adding low frequency switches which aids in doubling V_{RMS} and voltage levels eliminating the DC source midpoint. But in the circuit the additional implanted switches must need to operate at peak voltages which restricts the operation of MLI in high voltage applications.

Hybrid and asymmetric MLI were introduced in [7]. In asymmetric circuits, the marks of DC voltages are inconsistent or alter progressively [8]. These such setups decrease the size, complexity and even the cost of the converter and work on the unwavering quality since less capacitors, semiconductors and other basic elements are utilized [9]. The hybrid converters are made out of various staggered topologies with inconsistent levels of DC sources and distinctive regulation procedures. With proper determination of switching, the cost of converter is essentially diminished along with the loss of seclusion and issues with switching cycles and control procedures.

The association of the H bridge converters in series yield out the CHB circuits, which bring forth appealing qualities [9–13], including modular construction, simple execution with simple control circuits, generating sinusoidal voltages with zero electromagnetic impedance or common mode voltages, etc. This choice of inverter strategy and switching is a legitimate for significant level of applications according to the perspective in modularity and control.

Normally, NPC-PWM inverters in three phase using switches in each arm has considerable amount of suppression in harmonics in comparison to fundamental full-bridge PWM-based inverters. However, this is quite opposite in the case of single-phase PWM inverters. But, as of now, the 1-ph inverters embrace full-bridge using approximate techniques of modulation so that the voltage at output (V_{o}) has two values: zero voltage level and positive supply of V_{DC} in the positive waveforms. Consequently, harmonics in the output are examined by carrier frequency. But, however, the harmonics are also diminished.

On the above foundations, this article presents 17-level inverter topology, wherein, four PWM inverters are cascaded [14, 15], giving 17 voltage levels. This is particularly like the CHB inverter setup just as in every H-connection, a clipping switch has to be added to work on the general symphonious profile of output waves and moreover, the harmonic check of proposed inverter arrangement is exceptionally decreased when contrasted with the customary one-stage CHB, diode-clamped type inverter for a similar output level in voltage.

As numerous researches were carried out by researchers and enthusiasts there were many modulation techniques [16–21] in inverter switching mechanisms. The proposed inverter topology is contrasted with the CHB with different techniques of modulation such as LS-PWM, PS-PWM and NLC-PWM, POD-LS-PWM, APOD-LS-PWM and the same techniques of modulation are implemented for the newly proposed MLI with reduced switch count for examinations on functional standards and switching capacities. Simulation and test results are produced to confirm the legitimacy of the proposed topology against CHB inverter.

2 Cascaded 17 Level H Bridge Inverter

The conventional topology of cascaded 17 level H Bridge MLI is presented in Fig. 1. Every one of the four cells contains a solitary stage ordinary H-connect, with single bidirectional switch associated with the center tap of two sources of DC. Following the proper order of switching in inverter yields 17 levels of voltage as follows:

$8V_{dc}, 7V_{dc}, 6V_{dc}, 5V_{dc}, 4V_{dc}, 3V_{dc}, 2V_{dc}, V_{dc}, 0, V_{dc}, -V_{dc}, -2V_{dc}, -3V_{dc}, -4V_{dc}, -5V_{dc}, -6V_{dc}, -7V_{dc}$ and $-8V_{dc}$.

The voltage levels obtained in inverter is calculated for n quantities is $(4n + 1)$ for five-level inverters. By the switching sequence implemented, any voltage level underneath the peak can be integrated by changing the profile of modulation. Extra switches S_A to S_D should be appropriately triggered considering about the heading of the load current. The sequence of switching embraced in the cascaded inverter is shown in Fig. 2.

Figure 3 depicts output voltage in 1 per unit (PU) for one cycle for different switching angles which are plotted in the figure clearly. For each fundamental cycle component, the cascaded MLI operates through 17 levels which are as follows in Fig. 2.

The modulation index M_a is decided by phase angle θ depends on,

$$M_a = \frac{A_m}{A_c(n-1)}$$

A_c: value of triangular carrier signals from peak-to-peak
A_m: rectified sinusoidal ref signal at peak value
n: levels of voltage generated for one half-cycle; here, $n = 8$.

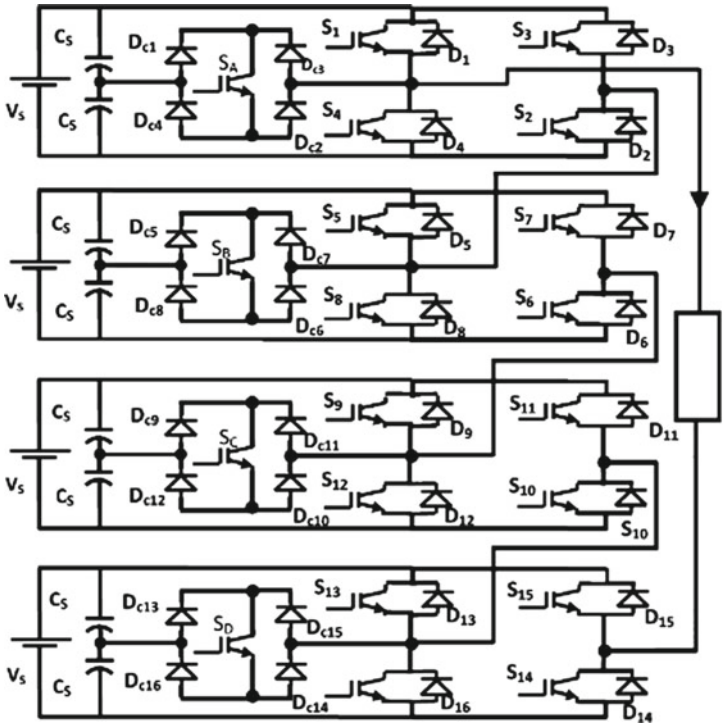
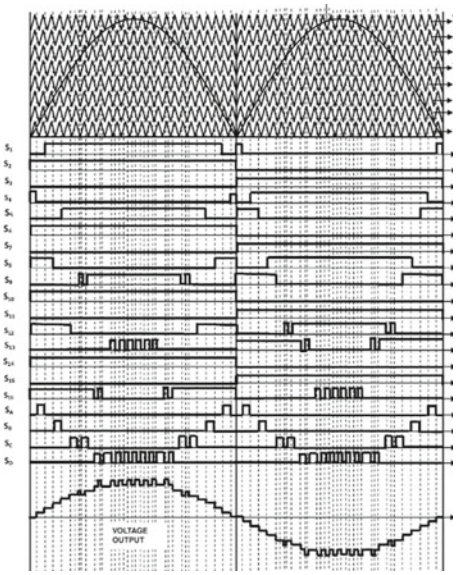


Fig. 1 17-level cascaded inverter



Level 1: $0 \leq \omega t \leq \theta_1$ & $\theta_{14} \leq \omega t \leq \pi$
Level 2: $\theta_1 \leq \omega t \leq \theta_2$ & $\theta_{13} \leq \omega t \leq \theta_{14}$
Level 3: $\theta_2 \leq \omega t \leq \theta_3$ & $\theta_{12} \leq \omega t \leq \theta_{13}$
Level 4: $\theta_3 \leq \omega t \leq \theta_4$ & $\theta_{11} \leq \omega t \leq \theta_{12}$
Level 5: $\theta_4 \leq \omega t \leq \theta_5$ & $\theta_{10} \leq \omega t \leq \theta_{11}$
Level 6: $\theta_5 \leq \omega t \leq \theta_6$ & $\theta_9 \leq \omega t \leq \theta_{10}$
Level 7: $\theta_6 \leq \omega t \leq \theta_7$ & $\theta_8 \leq \omega t \leq \theta_9$
Level 8: $\theta_7 \leq \omega t \leq \theta_8$
Level 9: $\theta = \pi$
Level 10: $\pi \leq \omega t \leq \theta_{15}$ & $\theta_{28} \leq \omega t \leq 2\pi$
Level 11: $\theta_{15} \leq \omega t \leq \theta_{16}$ & $\theta_{27} \leq \omega t \leq \theta_{28}$
Level 12: $\theta_{16} \leq \omega t \leq \theta_{17}$ & $\theta_{26} \leq \omega t \leq \theta_{27}$
Level 13: $\theta_{17} \leq \omega t \leq \theta_{18}$ & $\theta_{25} \leq \omega t \leq \theta_{26}$
Level 14: $\theta_{18} \leq \omega t \leq \theta_{19}$ & $\theta_{24} \leq \omega t \leq \theta_{25}$
Level 15: $\theta_{19} \leq \omega t \leq \theta_{20}$ & $\theta_{23} \leq \omega t \leq \theta_{24}$
Level 16: $\theta_{20} \leq \omega t \leq \theta_{21}$ & $\theta 22_8 \leq \omega t \leq \theta_{23}$
Level 17: $\theta_{21} \leq \omega t \leq \theta_{22}$

Fig. 2 Switching pattern embedded in the conventional 17 level MLI

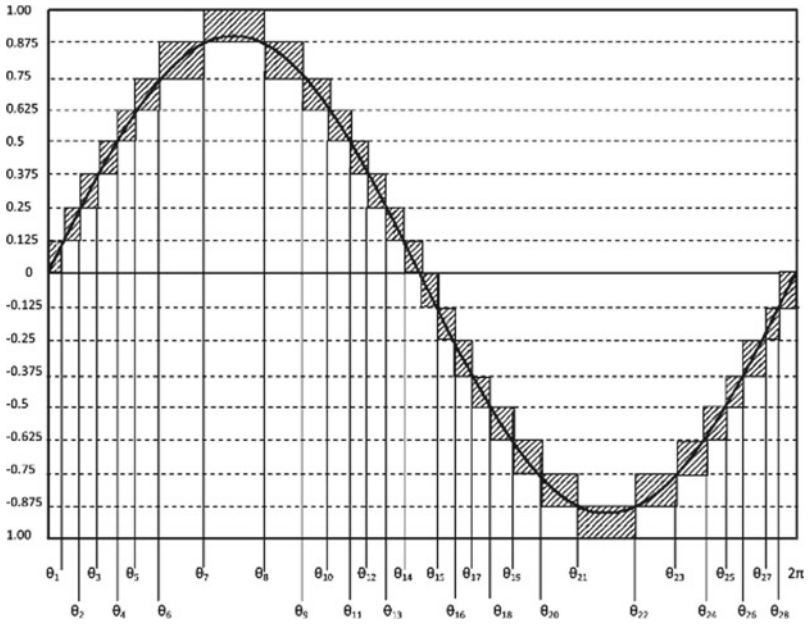


Fig. 3 17-level output voltage along with switching levels

The behavior of MLI has found similarities with conventional full-bridge 3 level PWM-based inverter [22]. The switching pattern exhibited by the different switches in the conventional topology is depicted in Table 1.

The phase disposition technique in PWM is utilized in gate signal generation. The essential standard of the switching technique is it has to produce gating signals by contrasting the readdressed sine signals at key recurrence, with the carrier waves which were having high switching freq (HZ) and also in phase yet with various offset voltages. Practically, the carrier to fundamental proportionate freq (HZ) at 40 is embraced for execution of topology [23].

3 Proposed MLI Configuration

In contrast to the previous conventional topology the proposed configuration with the reduced no' of switches in 17 level MLI is depicted in Fig. 4. This topology plots 4 asymmetric DC sources and 10 switches which generates 17 levels which are:

$8V_{dc}, 7V_{dc}, 6V_{dc}, 5V_{dc}, 4V_{dc}, 3V_{dc}, 2V_{dc}, V_{dc}, 0, V_{dc}, -V_{dc}, -2V_{dc}, -3V_{dc}, -4V_{dc}, -5V_{dc}, -6V_{dc}, -7V_{dc}$ and $-8V_{dc}$

which are same as that of in conventional cascaded H Bridge MLI discussed in Sect. 2. In comparison to topology Fig. 1 the proposed configuration in Fig. 4 eliminates the

Table 1 Switching sequence of switches in cascaded H bridged inverter

V_{load}	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	Sa	Sb	Sc	Sd
V_{dc}	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	On	Off	Off	Off
$2V_{dc}$	On	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	Off	Off	Off
$3V_{dc}$	On	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off
$4V_{dc}$	On	On	Off	Off	On	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	Off	Off	Off
$5V_{dc}$	On	On	Off	Off	On	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	Off	On	Off
$6V_{dc}$	On	On	Off	Off	On	On	Off	Off	On	On	Off	Off	Off	On	Off	Off	Off	Off	Off	Off
$7V_{dc}$	On	On	Off	Off	On	On	Off	Off	On	On	Off	Off	Off	On	Off	Off	Off	Off	Off	On
$8V_{dc}$	On	On	Off	Off	On	On	Off	Off	On	On	Off	Off	On	On	Off	Off	Off	Off	Off	Off
0	Off	On	Off	Off	Off	On	Off	Off	Off	Off	Off	Off	Off	Off	Off	Off	Off	Off	Off	Off
0	Off	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	Off	Off
$-V_{dc}$	Off	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off	On	Off	Off	Off
$-2V_{DC}$	Off	Off	On	On	Off	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off	Off	Off	Off	Off
$-3V_{dc}$	Off	Off	On	On	Off	Off	On	On	Off	Off	On	Off	Off	Off	On	Off	Off	On	Off	Off
$-4V_{dc}$	Off	Off	On	On	Off	Off	On	On	Off	Off	On	Off	Off	Off	On	Off	Off	Off	Off	Off
$-5V_{dc}$	Off	Off	On	On	Off	Off	On	On	Off	Off	On	Off	Off	Off	On	Off	Off	Off	On	Off
$-6V_{dc}$	Off	Off	On	On	Off	Off	On	On	Off	Off	On	On	Off	Off	On	Off	Off	Off	Off	Off
$-7V_{dc}$	Off	Off	On	On	Off	Off	On	On	Off	Off	On	On	Off	Off	On	Off	Off	Off	Off	On
$-8V_{dc}$	Off	Off	On	On	Off	Off	On	On	Off	Off	On	On	Off	Off	On	On	Off	Off	Off	Off

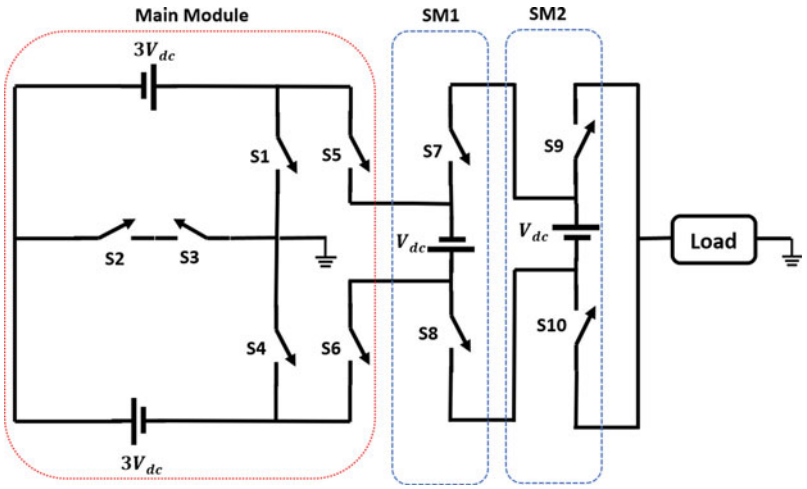


Fig. 4 Proposed configuration of 17 level

use of capacitors, diodes, and multi number of switches, making the design to be elegant and simple. The proposed configuration makes use of various dc source in generating the multi-levels of voltage in achieving 17 levels of output.

From Fig. 4 main module with two DC sources of $3V_{\mathrm{dc}}$ and 2 sub modules each comprising of one dc source of V_{dc} are associated in series through load to accomplish 17 level bidirectional voltage at load side.

Proposed plan and switching succession applied to switches will transform the unidirectional voltage into sinusoidal like bidirectional voltage. From the topology implemented, 10 switches are needed to generate 17 levels of voltage wave. Two switches introduced in each submodule are worked in free style to get bidirectional voltage at load. The execution of switching cycles of the proposed MLI configuration is cleared in Table 2.

4 Comparison of Works

The new reduced switch multi-level inverter topology is compared with the CHB MLI in various aspects and generalized equations are tabulated in Tables 3 and 4, respectively.

To learn the improvements of MLI arrangement, certain figures of with legitimacy are considered. A few of these are: required DC sources, number of orchestrated levels of voltage, number of dynamic and uninvolved switches, leading switches count needed to generate guaranteed voltage level, and so forth consequently, the details are to the size, nature of voltage generated and sifting level, etc. The proposed MLI displayed in Fig. 4 is contrasted with the conventional balanced CHB and some new

Table 2 Switching sequence of switches in proposed MLI

V_{load}	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
$8V_{dc}$	Off	Off	Off	On	On	Off	Off	On	On	Off
$7V_{dc}$	Off	Off	Off	On	On	Off	Off	On	Off	On
$6V_{dc}$	Off	Off	Off	On	On	Off	On	Off	On	Off
$5V_{dc}$	Off	On	On	Off	On	Off	Off	On	On	Off
$4V_{dc}$	Off	On	On	Off	On	Off	Off	On	Off	On
$3V_{dc}$	Off	On	On	Off	On	Off	On	Off	On	Off
$2V_{dc}$	On	Off	Off	Off	On	Off	Off	On	On	Off
$1V_{dc}$	On	Off	Off	Off	On	Off	Off	On	Off	On
0	On	Off	Off	Off	On	Off	On	Off	On	Off
$-1V_{dc}$	Off	Off	Off	On	Off	On	On	Off	On	Off
$-2V_{dc}$	Off	Off	Off	On	Off	On	On	Off	Off	On
$-3V_{dc}$	Off	On	On	Off	Off	On	Off	On	Off	On
$-4V_{dc}$	Off	On	On	Off	Off	On	On	Off	On	Off
$-5V_{dc}$	Off	On	On	Off	Off	On	On	Off	Off	On
$-6V_{dc}$	On	Off	Off	Off	Off	On	Off	On	Off	On
$-7V_{dc}$	On	Off	Off	Off	Off	On	On	Off	On	Off
$-8V_{dc}$	On	Off	Off	Off	Off	On	On	Off	Off	On

Table 3 Comparison of the new 17 level inverter with existing multi-level inverter

	CHB	Diode-clamped [30]	FCs [31]	Proposed MLI
Main switching device	20	32	32	10
Main diodes	0	56	0	0
DC bus	4	1	1	4
Balancing capacitor	0	0	56	0

Table 4 Comparison of the new MLI of 17-levels with CHB on circuit parameters basis

	CHB	Proposed MLI
DC sources used	$(n-1)/2$	$\frac{1}{6}(n+7)$
No. of output voltage level	N	N
Max. voltage generate	nV_s	nV_s
Peak inverse voltage	$4V_{ns}$	$(11nV_s)/3$
On-state switches	$2(n-1)$	$\frac{1}{3}(n+13)$

n = number of levels in MLI

PWM-based MLI techniques. The correlation, displayed in Tables 3 and 4, depends on per stage design of given number, n, with necessary info of DC source, V_S, in every one of the MLI topology.

For a predetermined number of voltage levels generated (V_o), the proposed circuit lessens the number of segregated dc sources by 32% in both CHB and also in topologies [24–28]; and furthermore by half in the setup introduced in [24]. Likewise, there is a decrease of 62.5% in count of dynamic switches in the presented inverter setup which when contrasted with the CHB and circuits in [24, 29]. This dynamic switch reduction is reflected by 40% in the configurations of [24–28]. The driver circuits and DC sources are also less in number in presented topology compared to other work. Overall the proposed 17 level inverter has proved to be efficient in many aspects in terms of number of components utilized, circuit parameters, harmonic content briefed in Sect. 5 in comparison to CHB and also the recent previous MLI techniques.

5 Simulation Results

To get to the legitimacy of the proposed inverter strategy, circuit simulations are worked out on MATLAB in accordance to the switching scheme presented in Tables 1 and 2, i.e., CHB topology and proposed topology against different modulation techniques of LS-PWM, PS-PWM and with half height NLC-PWM, respectively. Subsequently, by comparing the simulations are produced by PWM gating signals for switches S1 through S16 and SA through SD and gating signals from S1 to S10 in CHB and proposed MLI. All the simulations of output for voltage and current are carried out for the modulation index of $M_a = 1$.

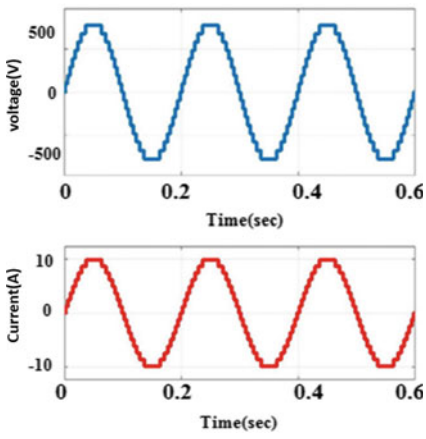


Fig. 5 V_o and I_L of CHB with LS-PWM technique

The simulations of CHB inverter with LS-PWM, PS-PWM, NLC-PWM, POD-PWM, APOD-PWM modulation strategies (Figs. 5, 6, 7, 8, 9, 10, 11, 12, 13 and 14).

The simulations of proposed reduced switch count MLI with LS-PWM, PS-PWM, NLC-PWM, POD-PWM, APOD-PWM modulation strategies (Figs. 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 and 25).

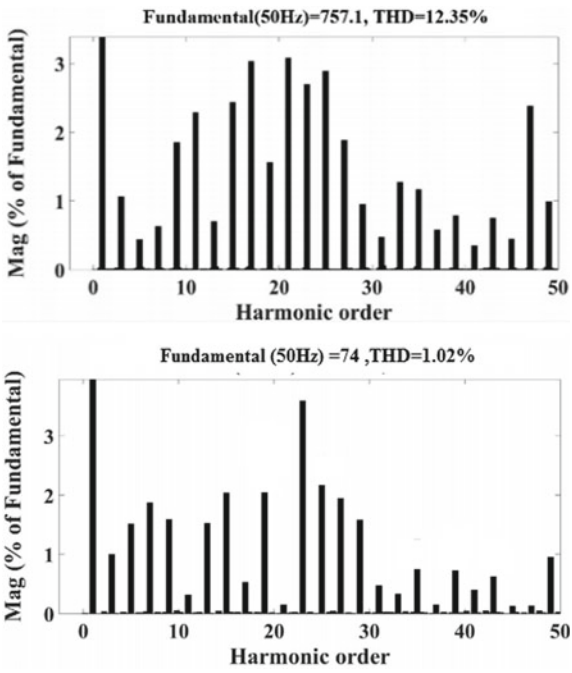


Fig. 6 THD of V_o and I_L of 17 level CHB with LS-PWM technique

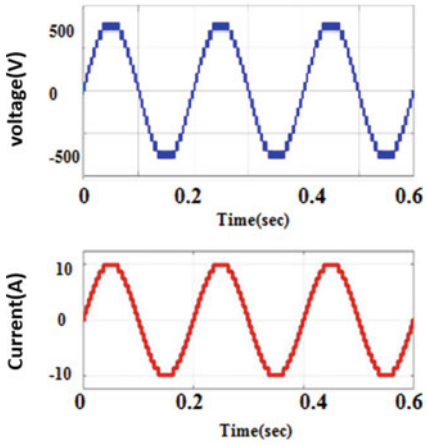


Fig. 7 V_o and I_L of 17 level CHB with PS-PWM

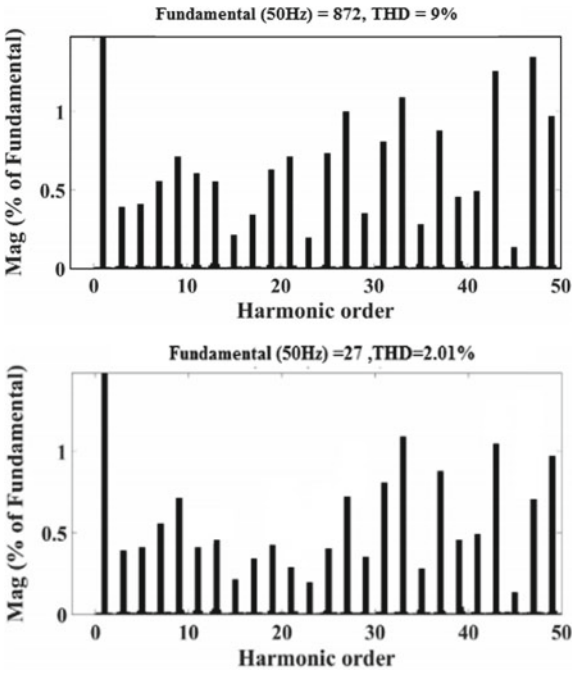


Fig. 8 THD of V_o and I_L of 17 level CHB with PS-PWM

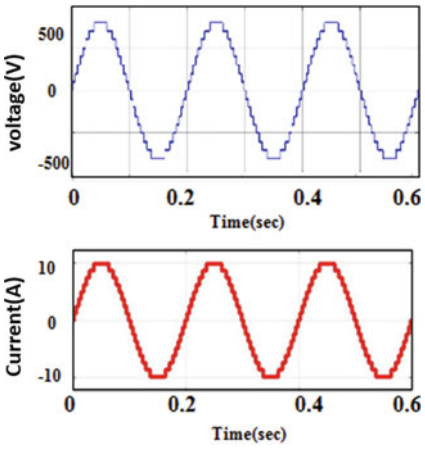


Fig. 9 V_o and I_L of 17 level CHB with NLC-PWM

The comparision of (%) THDs of outputs V_o and I_o in CHB and proposed reduced switch count MLI with different modulation schemes are presented in Table 5.

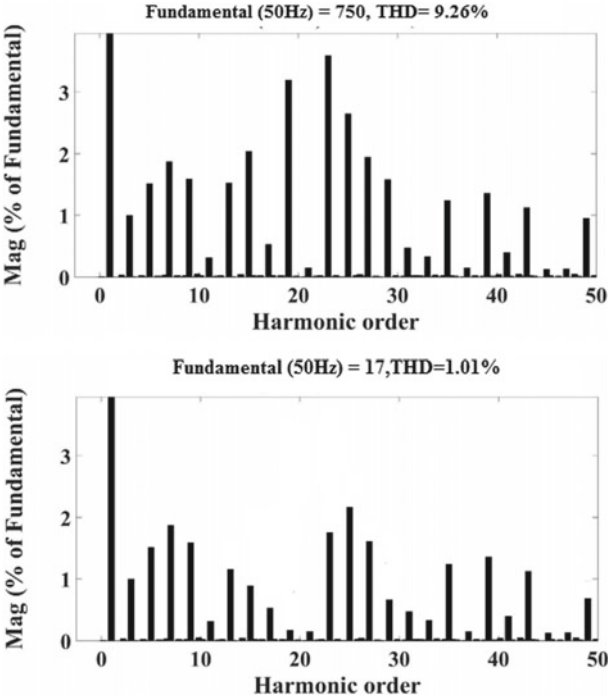
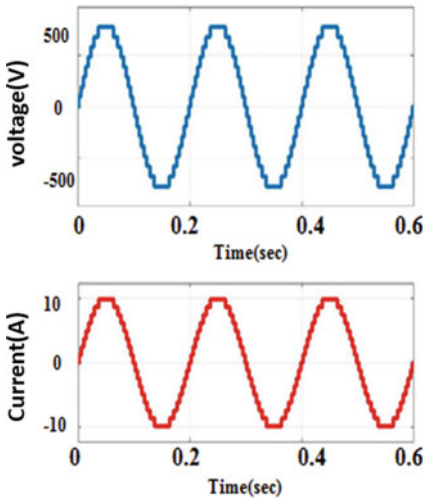


Fig. 10 THD of V_o and I_L of 17 level CHB with NLC-PWM

Fig. 11 V_o and I_L of 17 level CHB with POD-PWM



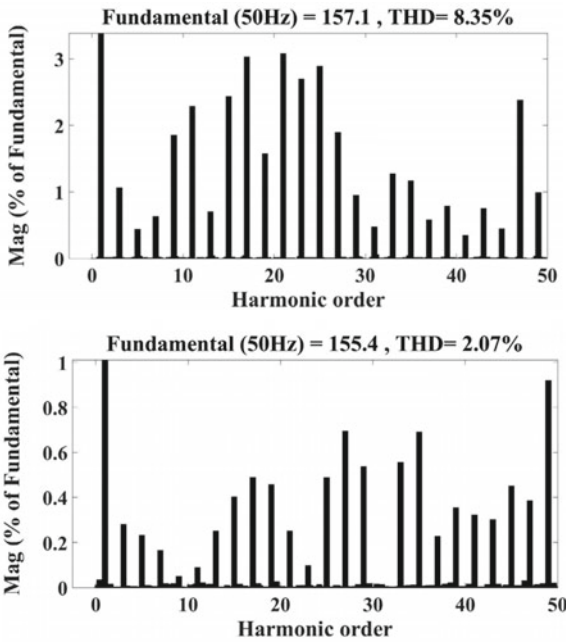


Fig. 12 THD of V_o and I_L of 17 level CHB with POD-PWM

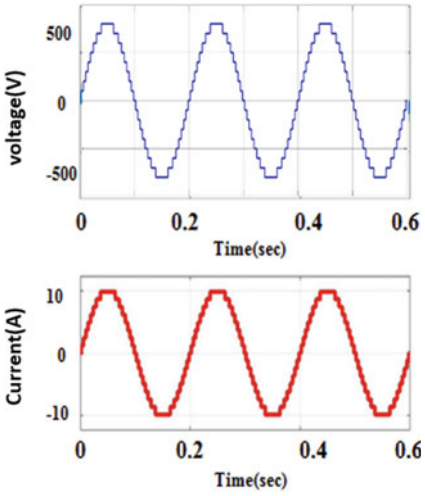


Fig. 13 V_o and I_L of 17 level CHB with APOD-PWM

6 Conclusion

This paper has proposed a reduced switch count MLI 17 level inverter in contrast to a CHB topology. The contrast work is carried out against various PWM modulation techniques such as with LS-PWM (level shift), PS-PWM (phase shift), half height

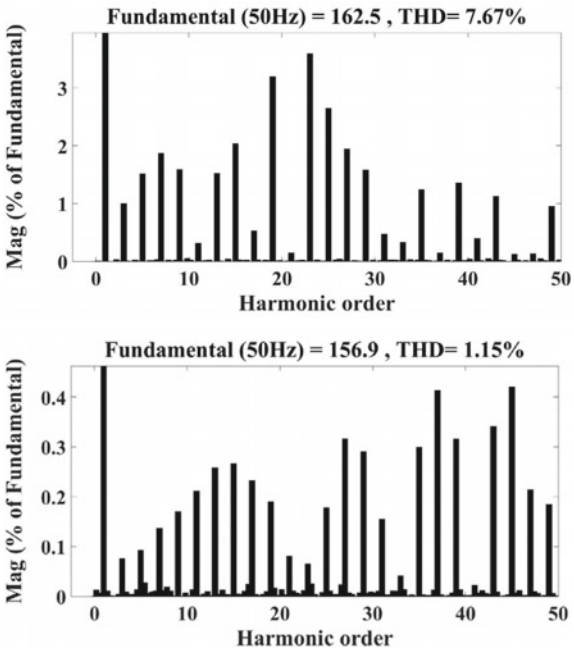


Fig. 14 THD of V_o and I_L of 17 level CHB with APOD-PWM

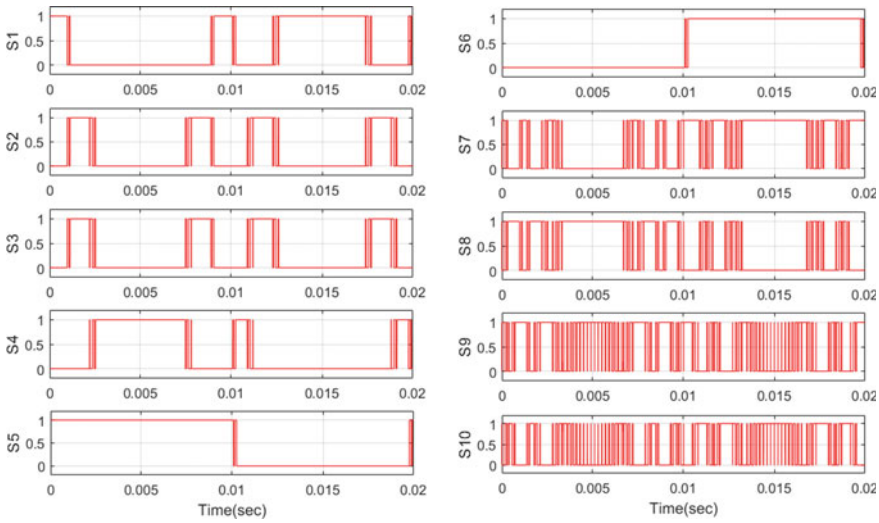


Fig. 15 Switching pulses of S1–S10 of proposed MLI with LS-PWM technique

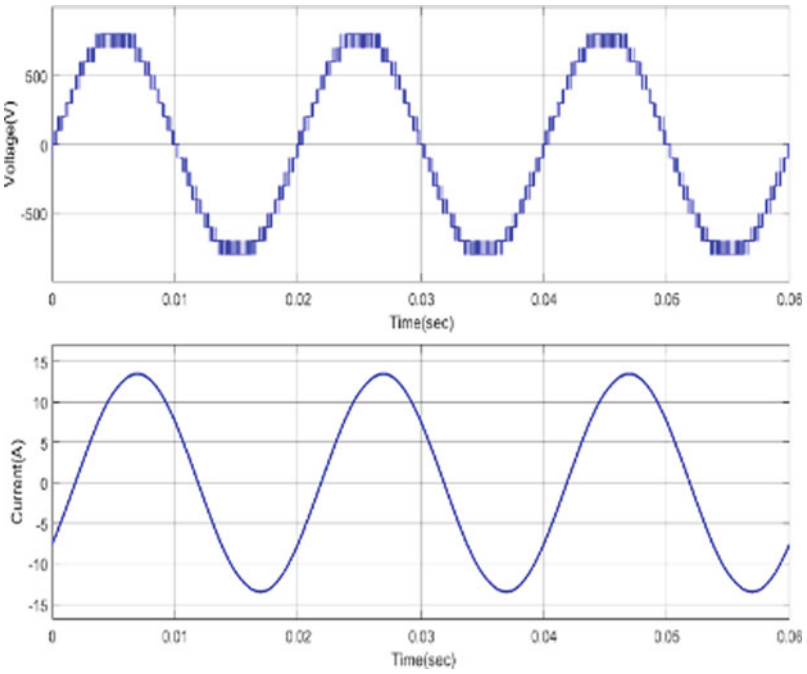


Fig. 16 V_o and I_L of 17 level MLI with LS-PWM technique

NLC-PWM (nearest level control). The simulation are carried out on the MATLAB and same simulations are presented. For assessing and validating the effectiveness of the proposed scheme, various MLI elements are contrasted in Tables 3 and 4 against various PWM modulation schemes. From the analysis carried out on various configurations its circuit parameters, it is clear that the output voltage and current THD has diminished to a great level in proposed MLI in contrast to CHB. On the other hand, the V_{rms}, I_{rms}, and power output are also found to be increased which has a great potential use to integrate within renewable energy sources (RES) [30, 31].

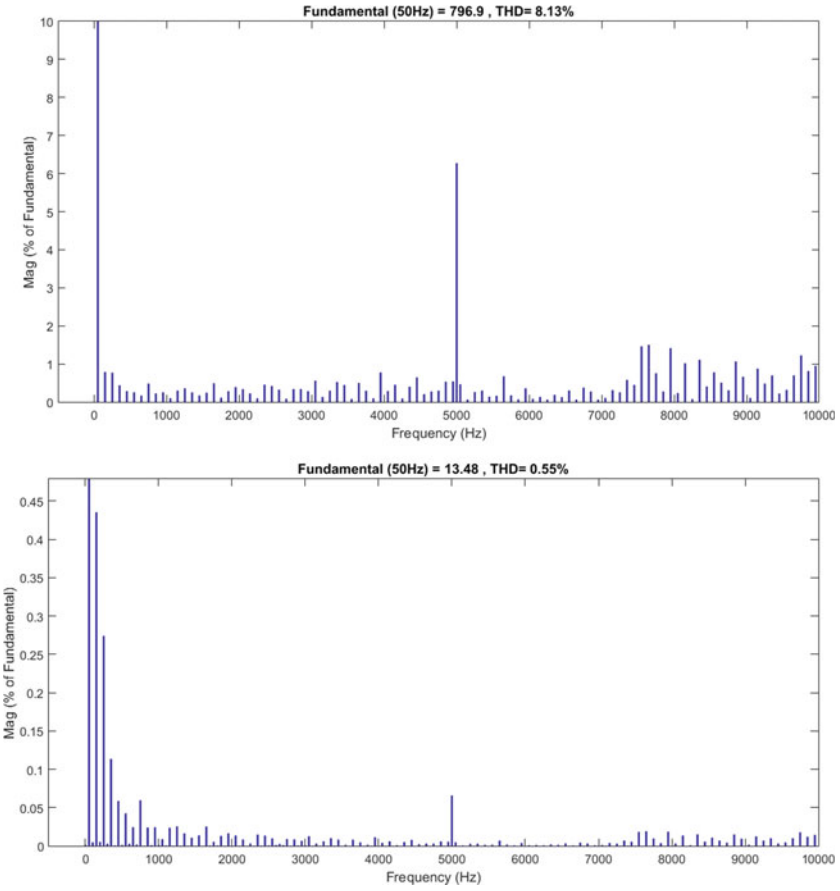


Fig. 17 THD of V_o and I_L of proposed MLI with LS-PWM technique

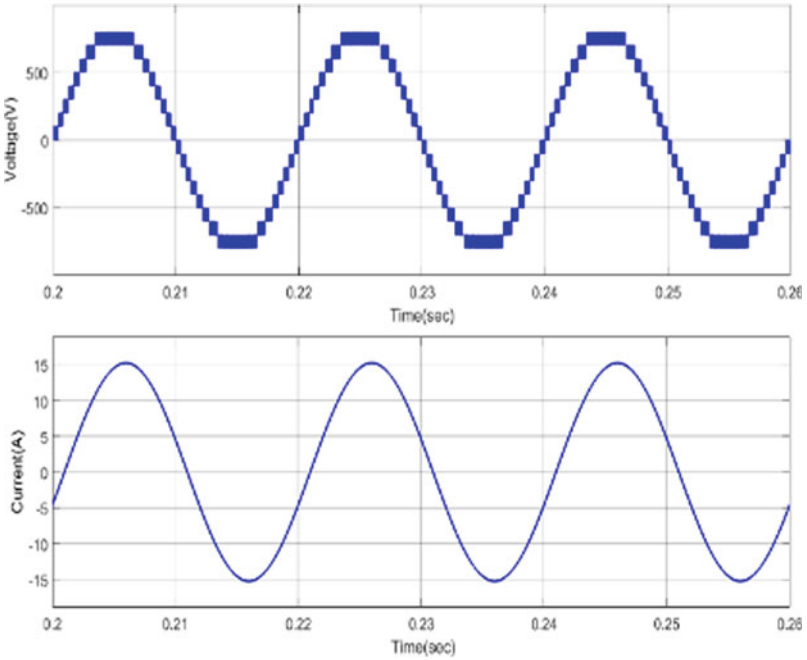


Fig. 18 V_o and I_L of proposed MLI with PS-PWM

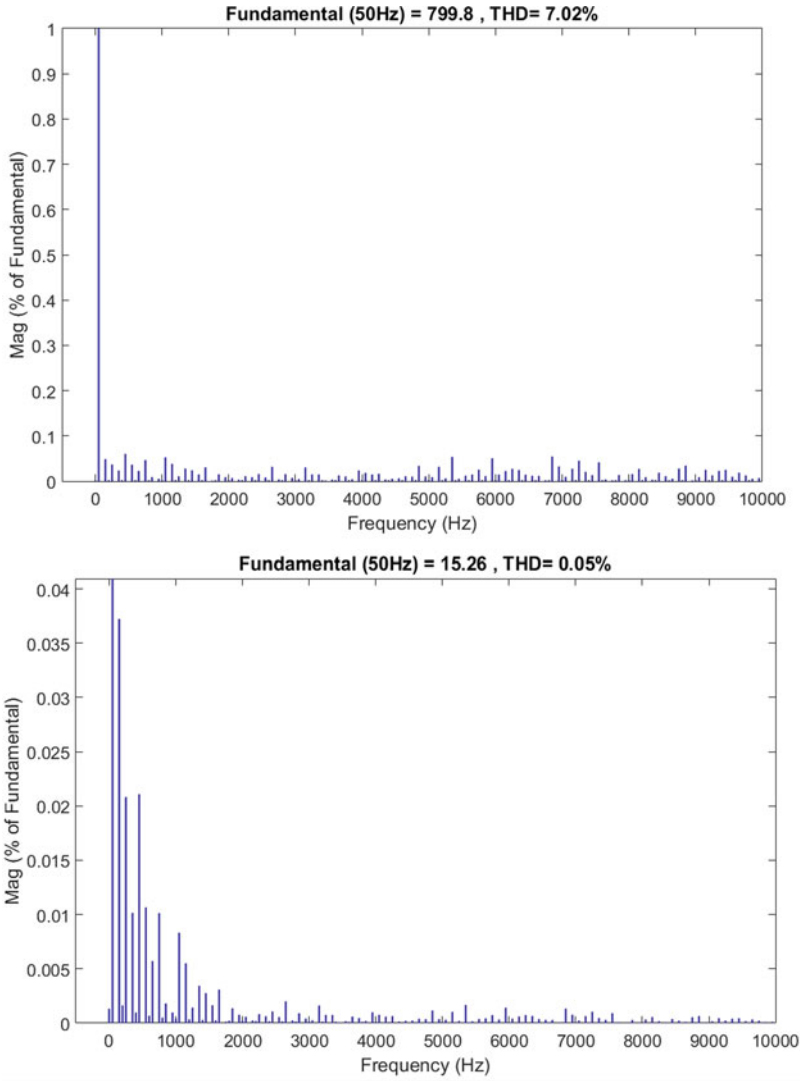


Fig. 19 THD of V_o and I_L of proposed MLI with PS-PWM

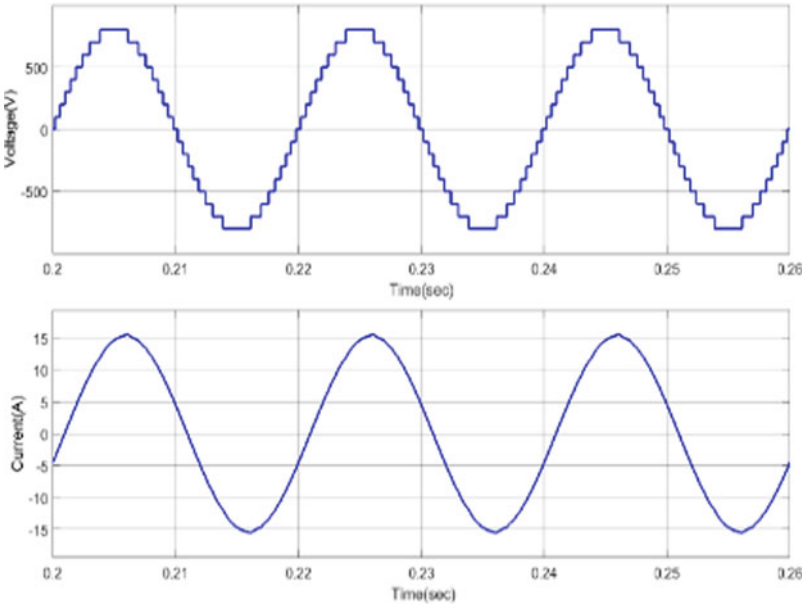


Fig. 20 V_o and I_L of proposed MLI with half height NLC-PWM

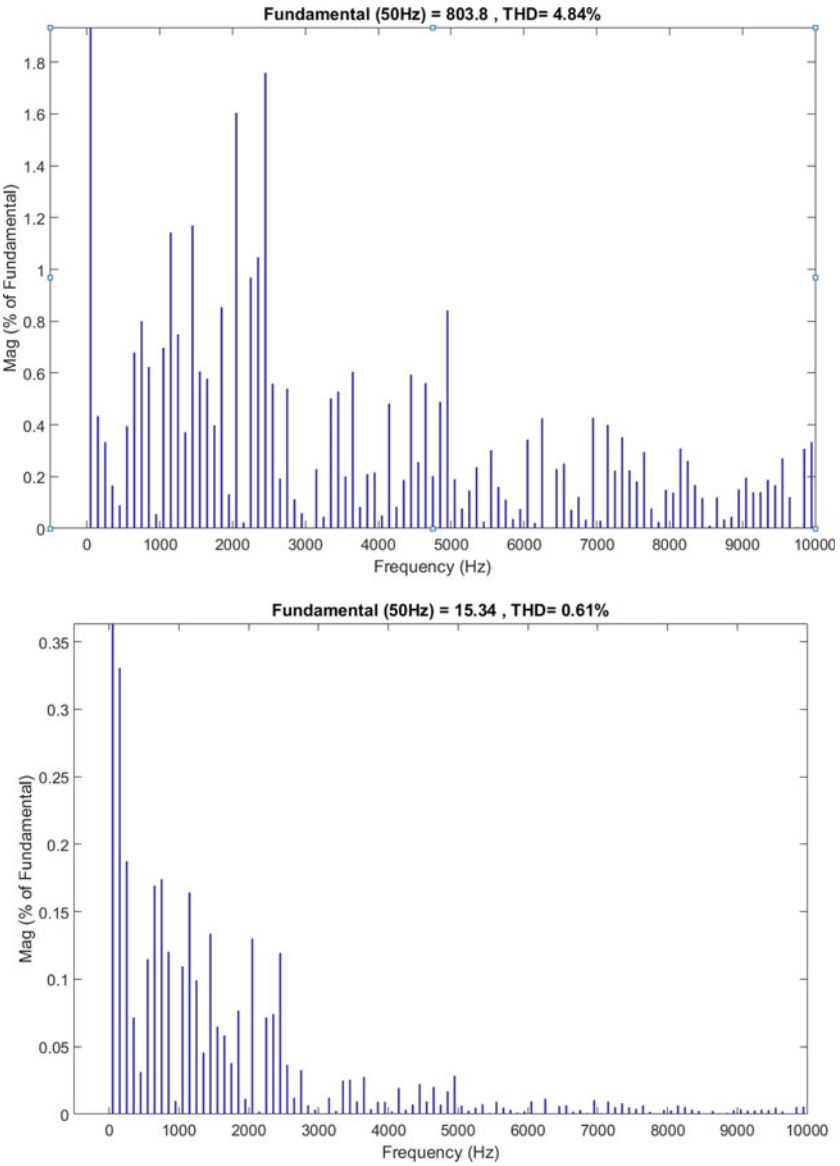


Fig. 21 THD of V_o and I_L of proposed MLI with NLC-PWM

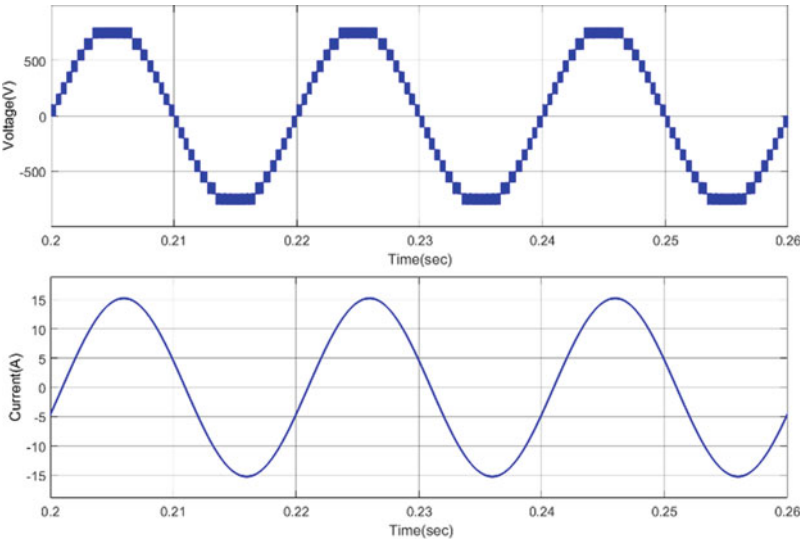


Fig. 22 V_o and I_L of proposed MLI with POD-LS-PWM

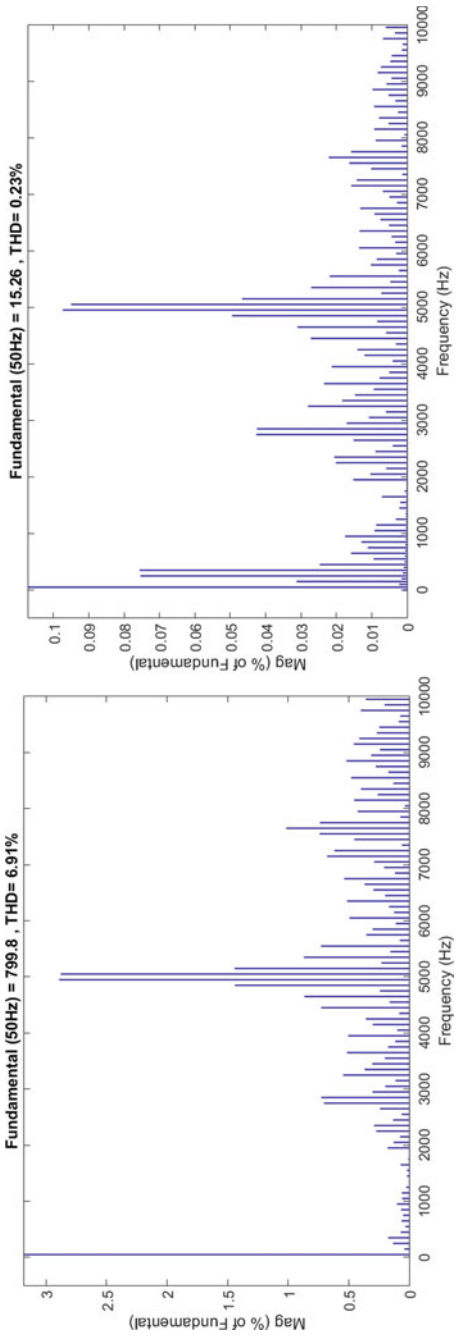


Fig. 23 THD of V_o and I_L of proposed MLI with POD-LS-PWM

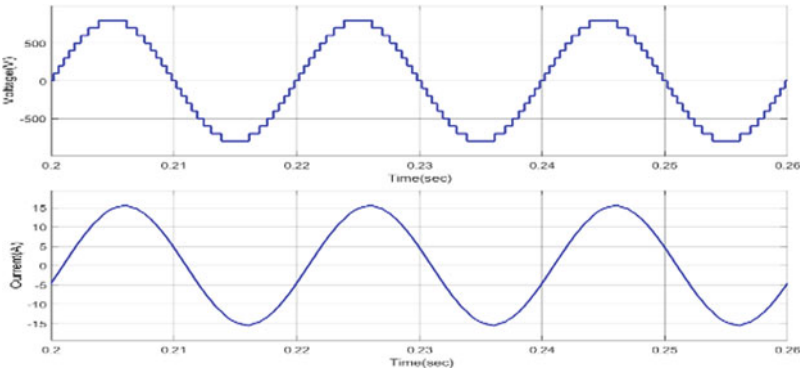


Fig. 24 V_o and I_L of proposed MLI with APOD-LS-PWM

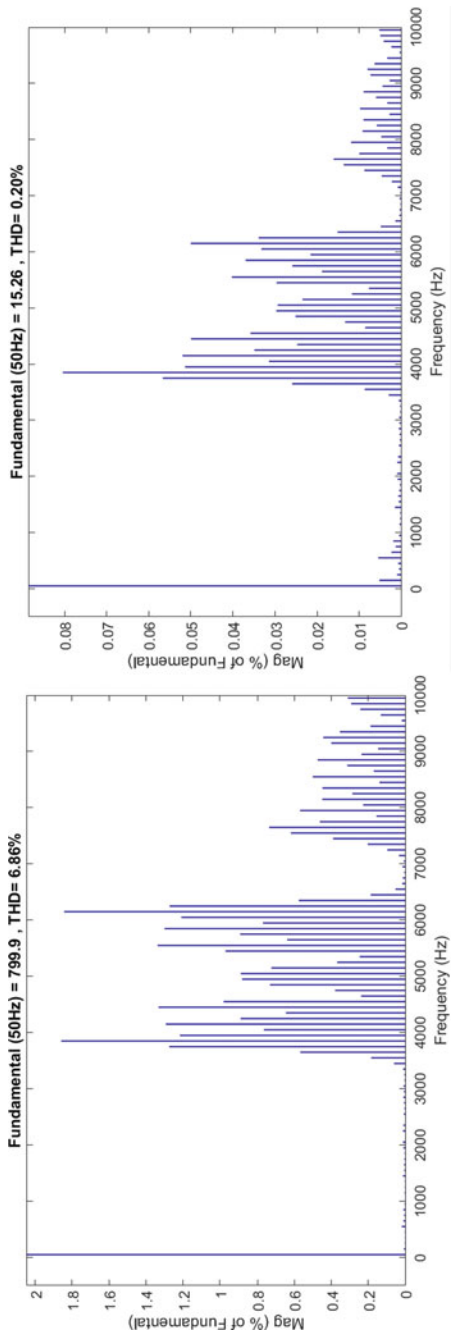


Fig. 25 THD of V_o and I_L of proposed MLI with APOD-LS-PWM

Table 5 Comparision of (%) THDs of output (a) voltage and (b) current

PWM schemes	(%) THD levels in voltage	
	CHB	Proposed MLI
(a) Voltage		
LS-PWM	12.35	8.13
PS-PWM	9	7.02
NLC-PWM	9.26	4.84
POD-PWM	8.35	6.91
APOD-PWM	7.65	6.86
PWM schemes	(%) THD levels in current	
	CHB	Proposed MLI
(b) Current		
LS-PWM	1.02	0.55
PS-PWM	2.01	0.05
NLC-PWM	1.01	0.61
POD-PWM	2.07	0.23
APOD-PWM	1.15	0.20

References

1. Nabae A, Takahashi I, Akagi H (1981) A new neutral-point-clamped PWM inverter. IEEE Trans Ind Appl 17(5):518–523
2. Choi NS, Cho JG, Cho GH (1991) A general circuit topology of multilevel inverter. In: Record of the 1991 IEEE power electronics specialists conference (PESC), Cambridge, MA, June 1991, pp 96–103
3. Meynard TA, Foch H (1992) Multi-level conversion: high-voltage chopper and voltage source inverters. In: Conference record of the 1992 IEEE power electronics specialists conference (PESC), Toledo, 29 June–3 July 1992, pp 397–403
4. Menzies RW, Steimer P, Steike JK (1993) Five-level GTO inverters for large induction motor drives. In: Conference record of the 1993 IEEE industry applications society (IAS) annual meeting, Oxford, UK, 2–6 May 1993, pp 595–601
5. Marchesoni M, Mazzucchelli M (1993) Multilevel converters for high power AC drives: a review. In: Proceedings of IEEE international symposium on industrial electronics (ISIE'93), Budapest, 1–3 June 1993, pp 38–43
6. Sadigh AK, Hosseini SH, Sabahi M, Gharehpetian GB (2010) Double flying capacitor multicell converter based on modified phase-shifted PWM. IEEE Trans Power Electron 25(6):1517–1526
7. Rech C, Pinheiro JR (2005) Line current harmonics reduction in multipulse connection of asymmetrically loaded rectifiers. IEEE Trans Ind Electron 52(3):640–652
8. Lu S, Marieethoz S, Corzine KA (2010) Asymmetrical cascade multilevel converters with noninteger or dynamically changing DC voltage ratios: concepts and modulation techniques. IEEE Trans Ind Electron 57(7):2411–2418
9. Rodriguez J, Franquelo LG, Kouro S, Leon JI, Portillo RC, Prats MAM, Perez MA (2009) Multilevel converters: an enabling technology for high-power applications. Proc IEEE 97(11):1786–1817
10. Malinowski M, Gopakumar K, Rodriguez J, Perez M (2010) A survey on cascaded multilevel inverters. IEEE Trans Ind Electron 57(7):2197–2206

11. Peng FZ, Lai J-S, McKeever JW, Van Coevering J (1996) A multilevel voltage-source inverter with separate DC sources for static VAr generation. IEEE Trans Ind Appl 32(5):1130–1138
12. Peng F, Lai JS (1997) Multilevel cascade VSI with separate DC sources. U.S. patent 5 642 275, 24 June 1997
13. Lezana P, Rodriguez J, Oyarzun JA (2008) Cascaded multilevel inverter with regeneration capability and reduced number of switches. IEEE Trans Ind Electron 55(3):1059–1066
14. Sridhar V, Umashankar S, Sanjeevikumar P, Ramachandaramurthy VK, Fedák V (2018) Control architecture for cascaded H-bridge inverters in large-scale PV systems. Energy Procedia 145:549–557
15. Gaikwad A, Arbune PA (2016) Study of cascaded H-bridge multilevel inverter. In: 2016 international conference on automatic control and dynamic optimization techniques (ICACDOT), pp 179–182. https://doi.org/10.1109/ICACDOT.2016.7877574
16. Lakshmi TVS, George N, Kothari DP (2013) Cascaded seven level inverter with reduced number of switches using level shifting PWM. In: Proceedings of 2013 international conference on power, energy and control (ICPEC), Feb 2013, pp 676–680
17. Edpuganti A, Rathore AK (2015) A survey of low switching frequency modulation techniques for medium-voltage multilevel converters. IEEE Trans Ind Appl 51(5):4212–4228
18. Hu P, Jiang D (2015) A level-increased nearest level modulation method for modular multilevel converters. IEEE Trans Power Electron 30(4):1836–1842
19. Prabaharan N, Palanisamy K (2017) A comprehensive review on reduced switch multilevel inverter topologies, modulation techniques and applications. Renew Sustain Energy Rev 76:1248–1282
20. Zaid MM, Ro J-S (2019) Switch ladder modified H-bridge multilevel inverter with novel pulse width modulation technique. IEEE Access 7:102073–102086
21. Grahame Holmes D, Lipo TA (2003) Carrier-based PWM of multilevel inverters. In: Pulse width modulation for power converters, principles and practice. Wiley, pp 453–526
22. Adam GP, Finney SJ, Massoud AM, Williams BW (2008) Capacitor balance issues of the diode-clamped multilevel inverter operated in a quasi two-state mode. IEEE Trans Ind Electron 55(8):3088–3099. https://doi.org/10.1109/TIE.2008.922607
23. Roshankumar P, Kaarthic RS, Gupakumar K, Leon JI, Franquelo LG (2015) A seventeen-level inverter formed by cascading flying capacitor and floating capacitor H-bridge. IEEE Trans Power Electron 30(7):3471–3478
24. Liu Z, Zhang Y, Zhao S, Gong J (2019) A power distribution control strategy between energy storage elements and capacitors for cascaded MLI with hybrid energy sources. IEEE Access 7:16880–16891
25. Alishah RS, Nazarpour D, Hosseini SH, Sabahi M (2014) Novel multilevel inverter topologies for medium and high voltage applications with lower values of blocked voltage by switches. IET Power Electron 7(12):3062–3071
26. Ajami A, Toupchi Khosroshahi M, Jannati Oskuee MR, Mokhberdoran A (2014) Cascade-multi-cell multilevel converter with reduced number of switches. IET Power Electr 7(3):552–558
27. Odeh CI, Nnadi DB (2013) Single-phase, 17-level hybridized cascaded multi-level inverter. Electr Power Compon Syst 41(2):182–196
28. Susheela N, Satish Kumar P (2017) Performance evaluation of multicarrier based techniques for single phase hybrid multilevel inverter using reduced switches. Indones J Electr Eng Comput Sci 7(3):676–686
29. Ajami A, Oskuee MRJ, Mokhberdoran A, Van den Bossche A (2014) Developed cascade-multi-cell multilevel inverter topology to minimise the number of circuit devices and voltage stresses of switches. IET Power Electr 7(2):459–466

30. Kiran R, Bhaskar MS, Padmanaban S, Blaabjerg F, Wheeler P, Rashid MH (2018) Unipolar single reference multicarrier sinusoidal pulse width modulation based 7-level inverter with reduced number of semiconductor switches for renewable energy applications. In: Proceedings of IEEE 18th international power electronics and motion control conference (PEMC), Aug 2018, pp 1043–1048
31. Lahlou T, Muresan VI, Bolvashenkov I, Herzog HG (2018) Space vector modulation for 17-level cascaded H-bridge inverter for use in battery energy storage systems. In: 2018 thirteenth international conference on ecological vehicles and renewable energies (EVER), pp 1–6. https://doi.org/10.1109/EVER.2018.8362379

A Review of Various Line Segmentation Techniques Used in Handwritten Character Recognition

Solley Joseph and Jossy George

Abstract Segmentation is a very critical stage in the character recognition process as the performance of any character recognition system depends heavily on the accuracy of segmentation. Although segmentation is a well-researched area, segmentation of handwritten text is still difficult owing to several factors like skewed and overlapping lines, the presence of touching, broken and degraded characters, and variations in writing styles. Therefore, researchers in this area are working continuously to develop new techniques for the efficient segmentation and recognition of characters. In the character recognition process, segmentation can be implemented at the line, word, and character level. Text line segmentation is the first step in the text/character recognition process. The line segmentation methods used in the character recognition of handwritten documents are presented in this paper. The various levels of segmentation which include line, word, and character segmentation are discussed with a focus on line segmentation.

Keywords Segmentation · OCR · Pattern recognition · Line segmentation · Word segmentation · Character segmentation · Touching lines segmentation · Overlapping lines segmentation

1 Introduction

1.1 Various Stages in the Character Recognition Process

Automatic recognition of printed or handwritten text from the images of the text documents and the conversion of it into a computer-readable form is called optical character recognition (OCR) [1]. The character recognition process comprises four

S. Joseph (✉) · J. George
CHRIST (Deemed to be University), Bengaluru, India
e-mail: solley.joseph@res.christuniversity.in

S. Joseph
Carmel College of Arts, Science and Commerce for Women, Goa, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_34

significant stages. The first stage is pre-processing, in which the document image is prepared for the process with the help of various techniques for noise removal and skew detection. Segmentation is the second stage, which involves the subdivision of the document image into components like words, lines, and characters. The next stage is the feature extraction, during which the distinguishing feature of each character is extracted. Classification is the fourth stage in which the input image is matched against a predefined set of classes.

1.2 Importance of Segmentation in OCR

Segmentation is an operation that subdivides the document image of a sequence of text into smaller components such as lines, words, and characters. It is one of the major steps in the process of optical character recognition (OCR) [2, 3]. The efficiency of segmentation has an impact on the performance of the entire character recognition process, which makes it a key stage in the character recognition process [4, 5]. This paper discusses the line, word, and character segmentation of handwritten documents with a focus line segmentation. The challenges in the segmentation of handwritten characters are also discussed. Segmentation has its applications in various fields in addition to character recognition. It is used in medical imaging, computer vision, object detection, and content-based image retrieval [6].

In the case of handwritten scripts, many factors such as overlapping or touching lines and characters, inconsistency in writing style, etc., influence the performance of the segmentation process [3, 7]. These challenges are discussed in Sect. 3.1. The nature of these challenges varies and depends on the type of script.

The rest of the sections are structured as follows—the various levels of text segmentation are given in Sect. 2. Sect. 3 explains various line segmentation methods, and Sect. 4 is the conclusion.

2 Levels of Segmentation

In the segmentation stage of the character recognition process, the image is broken down into lines, words, and characters. Although all these levels may be applied in a traditional OCR, it is not compulsory to use all these three levels. It depends on the nature of the application [8].

Character segmentation divides an image of a text string into sub-pictures of distinct characters. A character is a pattern that resembles one of the symbols, and the system is designed to recognize. During the character recognition process, the segmentation of characters has a critical function to perform, as the efficiency of character segmentation affects the character recognition accuracy directly [9, 10]. A review of cursive character recognition of handwritten documents is presented by Aravinda and Prakash [11]. Thungamani and Kumar [12] have performed a survey of

various approaches in Kannada character segmentation. The segmentation process is more challenging if the image is broken, touching, overlapping, and degraded and if there are compound characters. A survey of character segmentation methods used in cursive handwritten documents is presented by Choudhary [13]. Techniques for character segmentation are discussed.

Word segmentation is the task of separating the group of characters that correspond to words. It includes vertical image scan, pixel by pixel, left to right, top to bottom [14]. It can be performed by computing vertical projection profiles if the words are well-spaced. The complexity of the segmentation task increases in the case of handwritten characters. If there is no termination symbol for a word, segmentation of words is a very complex task. MODI script (a script that was used for writing Marathi till 1950) is an example of such a script where there is no termination symbol for words [15, 16]. Several researchers have implemented word segmentation for various languages. Some of the word segmentation methods used for various datasets are discussed in this section. Kumar et al. [17] performed an experiment on ten handwritten documents of Gurumukhi script using white space and pitch method and achieved an accuracy of 99.57%. Bansal and Sharma [18] experimented with a contour tracing algorithm on 2148 words of Gurumukhi script, which contains overlapping words, and the overall segmentation accuracy was 72.6%. In another experiment conducted by Ryu et al. [19], word segmentation was proposed as a binary quadratic problem, and the applied structured SVM framework was proposed for handwritten dataset (ICDAR 2009/2013), and the accuracy was 92.82%.

Line segmentation is the initial step in text-based image segmentation. In this step, the document image is scanned pixel by pixel, left to right, and top to bottom, and the image is then segmented into different regions with each one representing a text line [8, 14]. The focus for our study is line segmentation, and a detailed review of line segmentation methods is discussed in Sect. 3.

3 Line Segmentation

3.1 Challenges in Line Segmentation

Text line segmentation of handwritten documents faces many challenges. Most of the traditional segmentation algorithms work better in a situation where there is a distance between two adjacent text lines, and the text lines are almost straight. But in the case of handwritten text, small gaps between neighbouring text lines can result in overlapping lines and touching lines, and therefore, segmentation of lines of a handwritten text image is more complex compared to its printed counterparts. Slightly curved text lines may also be present in some handwritten documents. The presence of skewed lines is another factor that affects the line segmentation process. In the case of skewed lines, the skewed directions can also be different in some cases and that can aggravate the complexity of segmentation [20–23]. The complexity

increases in the case of historical and ancient manuscripts as they may be degraded and also contain overlapping or touching stamp symbols with the line segments. The quality of the scanned image is another factor that complicates the task. In short, the influence of the writer's style and the image quality affects the accuracy of the segmentation process [8, 24–26]. Researchers in the field of image processing have addressed several challenges of line segmentation and have developed various algorithms to increase the accuracy of line segmentation in such cases.

3.2 Segmentation Methodologies

The line segmentation methods can be categorized broadly into top-down and bottom-up approaches. In the top-down approach, the document is subdivided into text lines, words, and characters, recursively, in that order. This approach assumes that the lines are straight, and therefore, it works well on documents with reasonably straight text lines. Part of the images such as linked components, words, letters, and pixels are combined to generate a corresponding text line in the bottom-up technique. This approach is similar to the clustering process in which the text line is not assumed as a straight line. Hence, this approach is useful for documents containing lines with curved or overlapping text. The third category is a hybrid technique that blends both approaches in different ways [27, 28].

There are a number of strategies used for the bottom-up or top-down approaches, and they can be generally categorized into projection-based methods [17, 22, 27, 29–34], smearing [21, 25, 35–38], grouping methods [39–41], Hough-based methods [42–48], and graph-based methods [49–52].

Projection-based method

The most commonly used method in line segmentation is the projection-based method. This method calculates the vertical projection profile by adding the pixel values of each '*y*' value along the horizontal axis. The distance between the lines can be found from the vertical profile. Although this approach is more suitable for printed documents, some researchers modify this method and use a partial projection profile so that it can be used for handwritten documents with touching and overlapping lines as well.

Tripathy and Pal [22] used a projection-based method to perform the line segmentation of Oriya text by dividing the document into vertical stripes. The height of the stripe was calculated using the water reservoir method. The stripe-wise horizontal histogram is then calculated, and piecewise separating lines (PSL) are extracted from each of the stripes. Piecewise separating lines are joined to get a line segment using a two-step approach. A similar method of strip-based projection profile approach is implemented by Kumar et al. [17] for the line segmentation of documents containing handwritten Gurmukhi characters. The authors analyzed the projection profiles of foreground pixels in every strip as shown in Fig. 1.

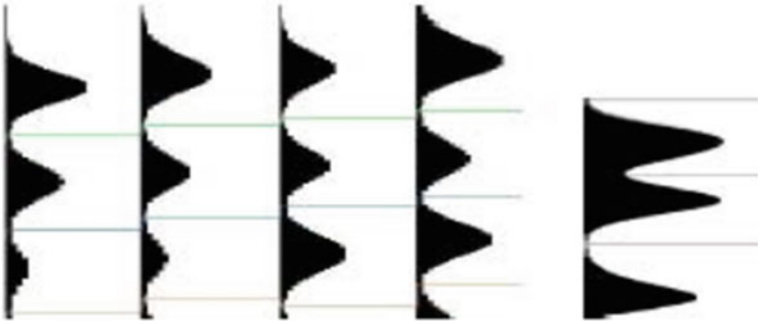


Fig. 1 Strip-based projection profile [17]

An algorithm based on oriented local projection (LPP) is implemented by Bar-Yosef et al. [29]. The experiment was carried out on manuscripts with multi-skew angles and with curved text lines, and very accurate results have been achieved. Timar and Karakas [30] have devised a method to localize the lines by calculating horizontal histograms for the whole image at a few pertinent skew angles. The angle and the position where the histograms have local minima are chosen as the location between lines. The piecewise projection profile approach is experimented by Garg and Garg [31] for the line segmentation of skewed and overlapped Devanagari script documents and it showed promising results. Arivazhagan et al. [32] used a similar technique to obtain the initial set of candidate lines. This method was reported to be efficient in dealing with documents containing skewed and touching lines. Figure 2 depicts the piece-projection profile experimented by Arivazhagan et al. [32].

a) The original document b) Piecewise projection profile

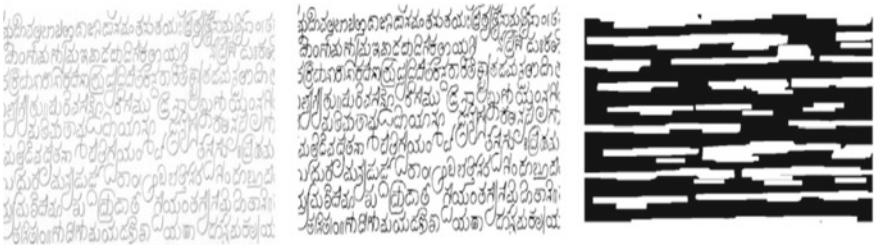
Fig. 2 An illustration of piecewise projection profile implementation [32]

Smearing method

Smearing techniques enhance the text area by smearing the consecutive black pixels along the horizontal direction. This is accomplished by filling the white region with black pixels in case, the distance between them is within a predefined threshold [17]. An analysis of the performance of major smearing techniques such as run-length smearing algorithm (RLSA), binary transition count map, and adaptive local connectivity map was carried out by Vishwas et al. [35]. The result of the study indicates that the binary transition count map technique is more effective compared to the run-length smoothing method in the case of documents with touching and overlapping lines. Roy et al. [36] implemented morphological operation and RLSA to extract the text lines from handwritten document images of Bangla, Devanagari, English, Oriya, and Gujarati scripts and have achieved encouraging results. Shi and Govindaraju [37] implemented a text line localization and separation algorithm by constructing a fuzzy run-length matrix. This matrix is used for segmenting the content of a manuscript into lines and words and other graphic areas. Figure 3 shows the various stages of the implementation of RLSA by Vishwas and Thomas [35].

Grouping method

This is a method in which the units such as pixels, and connected components are put together for constructing alignments. A method for text line extraction which is based on perceptual grouping based on connected components of black pixels is presented by Likforman-Sulem [41]. This method was able to detect multi-skewed documents by stepwise processing of the whole document from local to global. Line segments are created iteratively by combining adjacent linked components based on perceptual criteria such as resemblance, continuity, and proximity. As a result, local constraints on the neighbouring elements are blended with global quality measures. The technique uses a refining process that combines global and local analysis to handle conflicts. This method is not ideal for degraded or badly structured materials.



(a) Sample image (b) Binarized Image of script (c) Map obtained after applying RLSA

Fig. 3 Implementation of RLSA on Kannada script [35]

Fig. 4 Hough-based text segmentation [48]

Hough man-based method

The Hough transform is a popular algorithm that extracts the standard shape within an image [42]. This method is successfully implemented for line and word segmentation. Likforman-Sulem et al. [43] implemented a Hough-based method for extracting text lines in handwritten documents. A block-based Hough transform mapping was brought in by Louloudis et al. [44] for the detection of the line in handwritten documents. Figure 4 shows text segmentation from multi-script handwritten document images using Hough-based method, implemented by Saha et al. [48].

Graph-based method

Graph-based method is presented by Abuhaiba et al. [49], in which the lines are extracted by detecting the shortest spanning tree of a graph created from the collection of primary strokes. A graph-cut based method is implemented for the segmentation of Telugu documents scripts by Sesh Kumar et al. [50]. The authors propose that this strategy can be used for other Indian scripts as well; however, the traditional segmentation algorithm may not give good results in the case of complicated scripts, such as those used in Indian languages.

Segmentation using the stochastic method [21, 52] and cut text minimization method (CTM) [23, 53] is also implemented by various researchers (which are not included in the above-mentioned categories). In addition, there are other segmentation methods that cannot be included in any of the above-mentioned categories as they do not share common guidelines. A recent trend in line segmentation shows the use of neural networks and deep learning-based methods. Vo et al. [54] have carried out line segmentation on ICDAR2013 Handwritten Segmentation Contest dataset, using a fully convolutional network (FCN) in handwritten document images and achieved an accuracy of 98.68%. In this experiment, FCN is applied for building text string, and a line adjacency graph (LAG) is implemented to extract the text lines. A variant of deep fully convolutional network (FCN) is implemented by Renton et al. [55]. Using this method which is based on the dilated convolutions, they have conducted the experiment on two different data set; one is the ICDAR 2017 competition data set and the other one is a privately created data set. A consolidated list of some segmentation techniques that are most widely implemented by various researchers and their performance on the specific datasets are listed in Table 1.

Table 1 Performance comparison of line segmentation techniques

No	Category	Author	Method	Accuracy	Data set/language	Suitable for
1	Projection profile	(a) Tripathy and Pal [22]	Projection-based	97%	1627 text lines of Oriya script	Touching/skewed
		(b) Weliwitage et al. [23]	Cut text minimization	96%	NIST database	Distorted documents
		(c) Bar-Yosef et al. [29]	Oriented local projection profile	98%	30 degraded historical documents	Degraded documents with different skew angles
		(d) Arivazhagan et al. [32]	Bivariate Gaussian	Database 1—97.01% Database 2—98.62%	Database 1—CDARFOX Database 2—Arabic database	Skewed documents and overlapping lines
		(e) Garg and Garg [31]	Piecewise projection profiles	92.8%	Handwritten Devanagari script	Overlapping lines
		(f) Timar and Karakas [30]	Histogram—convex hull	87.06%	25 handwritten pages	Skewed lines
		(g) Kumar et al. [17]	Strip-based projection profile	93.7%	10 handwritten document of Gurumukhi script	Overlapping lines
		(h) Garg et al. [56]	Two stripe projection method	94.5%	200 lines of handwritten Devanagari script	Variable skew lines of text
		(i) Deshmukh et al. [57]	Thresholding histogram	98.41%	2540 MODI script documents	Multi-skew/touching/degraded
2	Smearing method	(a) Shi and Govindaraju [37]	Run-length smearing algorithm	93%	1864 USPS postal parcel images	Complex handwritten document

(continued)

Table 1 (continued)

No	Category	Author	Method	Accuracy	Data set/language	Suitable for
		(b) Roy et al. [36]	RLSA and morphological operations	Bangla—94.23% Devanagari—93.33% English—96.06% Oriya—94.76% Gujarati—93.04%	Multi-script documents	Unconstrained handwritten documents
3	Neural networks and deep learning-based	(a) Vo et al. [54]	Fully convolutional network	98.68%	350 documents of ICDAR 2013 handwritten segmentation test dataset	Parallel text lines
		(b) Renton et al. [55]	Deep convolutional network	93%	Dataset 1—dataset of 1600 images Dataset 2—from ICDAR competition	Handwritten documents
4	Morphological	Kulkarni et al. [58]	Morphological technique	94% (touching) 91% (overlapping)	50 historical document images	Historical handwritten documents

4 Conclusion

Segmentation is a major stage in the process of character recognition, and it has an impact on the accuracy of the system. Generally, the segmentation task is more complex in the case of handwritten text compared to its printed counterparts. The level of complexity in the segmentation of handwritten documents varies depends on many factors. A handwritten may contain lines that are skewed or overlapped, or it can have touched, overlapping, broken, or degraded characters. All these factors cause challenges in the segmentation process. In this paper, we have reviewed the line segmentation techniques used in handwritten documents by several researchers, in addressing these challenges. Following a brief discussion of the levels of segmentation, the line segmentation methodologies that are generally used are elaborated. A comparative study of the performance of various methods is also presented. This paper helps those who intend to carry out research in the area of line segmentation of handwritten text as they get guidance on the methodologies used in addressing specific complexities of line segmentation. The information provided in the tabular format helps the reader to get a quick review of all the methods which are implemented for addressing the specific type of dataset.

References

1. Joseph S, George JP, Gaikwad S (2020) Character recognition of MODI script using distance classifier algorithms. In: Fong S, Dey N, Joshi A (eds) ICT analysis and applications. Lecture notes in networks and systems, vol 93. Springer, Singapore
2. Casey RG, Casey RG, Lecolinet E, Lecolinet E (1996) A survey of methods and strategies in character segmentation. Analysis 18(7):690–706
3. Joseph S, George J (2021) Convolutional autoencoder based feature extraction and KNN classifier for handwritten MODI script character recognition. In: Shukla S, Unal A, Varghese Kureethara J, Mishra DK, Han DS (eds) Data science and security. Lecture notes in networks and systems, vol 290. Springer, Singapore
4. Joseph S, George J (2019) Feature extraction and classification techniques of MODI script character recognition. Pertanika J Sci Technol 27(4):1649–1669
5. Sahare P, Dhok SB (2019) Multilingual character segmentation and recognition schemes for Indian document images. IETE Tech Rev (Inst Electron Telecommun Eng India) 36(2):209–222
6. Jung K, Kim K, Jain A (2004) Text information extraction in images and video: a survey. Pattern Recognit 977–999
7. Joseph S, George J (2021) Data augmentation for handwritten character recognition of MODI script using deep learning method. In: Senjyu T, Mahalle PN, Perumal T, Joshi A (eds) Information and communication technology for intelligent systems. ICTIS 2020. Smart innovation, systems and technologies, vol 196. Springer, Singapore
8. Dos Santos RP, Clemente GS, Ren TI, Calvalcanti GDC (2009) Text line segmentation based on morphology and histogram projection. In: Proceedings of international conference on document analysis and recognition, ICDAR, May 2014, pp 651–655
9. Mathew CJ, Shinde RC, Patil CY (2011) Segmentation techniques for handwritten script recognition system. In: Signal processing, Aug 2011, pp 304–309
10. Joseph S, George J (2022) Offline character recognition of handwritten MODI script using wavelet transform and decision tree classifier

11. Aravinda CV, Prakash HN (2017) A review on various handwritten cursive characters segmentation techniques. In: Proceedings of 2016 2nd international conference on applied and theoretical computing and communication technology. iCATccT 2016, pp 647–651
12. Thungamani M, Kumar PR (2012) A survey of methods and strategies in handwritten Kannada character segmentation. Int J Sci Res 01(01):18–23
13. Choudhary A (2014) A review of various character segmentation techniques for cursive handwritten words recognition. Int J Inf Comput Technol 4(6):559–564
14. Dave N (2015) Segmentation methods for hand written character recognition. Int J Signal Process Image Process Pattern Recognit 8(4):155–164
15. Joseph S, Datta A, Anto O, Philip S, George J (2021) OCR system framework for MODI scripts using data augmentation and convolutional neural network. In: Data science and security. Lecture notes in networks and systems, vol 132. Springer, Singapore, pp 201–209
16. Joseph S, George J (2021) Efficient handwritten character recognition of MODI script using wavelet transform and SVD. In: Data science and security. Lecture notes in networks and systems, vol 132. Springer, Singapore
17. Kumar M, Sharma RK, Jindal MK (2010) Segmentation of lines and words in handwritten Gurmukhi script documents. In: ACM international conference proceeding series, pp 25–28
18. Bansal G, Sharma D (2010) Isolated handwritten words segmentation techniques in Gurmukhi script. Int J Comput Appl 1(24):122–130
19. Ryu J, Il Koo H, Cho NI (2015) Word segmentation method for handwritten documents based on structured learning. IEEE Signal Process Lett 22(8):1161–1165
20. Solley T (2018) A study of representation learning for handwritten numeral recognition of multilingual data set. In: Lecture notes in networks and systems, vol 10. Springer, pp 475–482
21. Brodic D (2010) A new approach to water flow algorithm for text line segmentation. Computer (Long Beach Calif) 17(1):30–47
22. Tripathy N, Pal U (2006) Handwriting segmentation of unconstrained Oriya text. Sadhana Acad Proc Eng Sci 31(6):755–769
23. Weliwitage C, Harvey AL, Jennings AB (2005) Handwritten document offline text line segmentation. In: Digital image computing: techniques and applications (DICTA'05), Jan 2006. IEEE, p 27
24. Momin A, Sanyal S (2012) Novel approach for segmenting fused/merged characters during character segmentation, pp 344–348
25. Likforman-Sulem L, Zahour A, Taconet B (2007) Text line segmentation of historical documents: a survey. Int J Doc Anal Recognit 9(2–4):123–138
26. Joseph S, George J (2020) Handwritten character recognition of MODI script using convolutional neural network based feature extraction method and support vector machine classifier. In: 2020 IEEE 5th international conference on signal and image processing (ICSIP), Nanjing, pp 32–36
27. Kumar MR (2012) Text line segmentation of handwritten documents using clustering method based on thresholding approach. Int J Comput Appl 975:8878
28. Bukhari SS, Shafait F, Breuel TM (2011) Improved document image segmentation algorithm using multiresolution morphology. Doc Recognit Retr XVIII 7874:78740D
29. Bar-Yosef I, Hagbi N, Kedem K, Dinstein I (2009) Line segmentation for degraded handwritten historical documents. In: Proceedings of international conference on document analysis and recognition, ICDAR, pp 1161–1165
30. Timar G, Karakas K (2003) Analogic preprocessing and segmentation algorithms for off-line handwriting recognition. J Circuits Syst Comput 12(06):783–804
31. Garg R, Garg NK (2014) An algorithm for text line segmentation in handwritten skewed and overlapped Devanagari script. Int J Emerg Technol Adv Eng 4(5):114–118
32. Arivazhagan M, Srinivasan H, Srihari S (2007) A statistical approach to line segmentation in handwritten documents. Doc Recognit Retr XIV 6500:65000T. International Society for Optics and Photonics
33. Bal A, Saha R (2016) An improved method for handwritten document analysis using segmentation, baseline recognition and writing pressure detection. Procedia Comput Sci 93:403–415

34. Manmatha R, Rothfeder JL (2005) A scale space approach for automatically segmenting words from historical handwritten documents. IEEE Trans Pattern Anal Mach Intell 27(8):1212–1225
35. Vishwas HS, Thomas BA (2015) Impact of smearing techniques on text line localization of Kannada historical scripts. Int J Comput Appl 975:8887
36. Roy P, Pal U, Lladós J (2008) Morphology based handwritten line segmentation using foreground and background information. In: International conference on frontiers in handwriting recognition, pp 241–246
37. Shi Z, Govindaraju V (2004) Line separation for complex document images using fuzzy runlength. In: First international workshop on document image analysis for libraries, June 2004. IEEE, pp 306–312
38. Shafait F, Keysers D, Breuel TM (2006) Performance comparison of six algorithms for page segmentation. In: International workshop on document analysis systems. Springer, Berlin, Heidelberg, pp 368–379
39. Feldbach M, Tönnies KD (2001) Line detection and segmentation in historical church registers. In: Proceedings of international conference on document analysis and recognition, ICDAR, Dec 2001, vol 2001, pp 743–747
40. Ben Messaoud I, Amiri H, El Abed H, Märgner V (2012) A multilevel text line segmentation framework for handwritten historical documents. In: Proceedings of international workshops on frontiers in handwriting recognition, IWFHR, pp 515–520
41. Likforman-Sulem L (1994) Extracting lines on handwritten documents by perceptual grouping. In: Advances in handwriting and drawing: a multidisciplinary approach, pp 117–135
42. Hough PVC (1962) Methods and means for recognizing complex patterns
43. Likforman-Sulem L, Faure C, Nationale E (1995) A Hough based algorithm for extracting text lines in handwritten documents, pp 774–777
44. Louloudis G, Gatos B, Pratikakis I, Halatsis K (2006) A block-based Hough transform mapping for text line detection in handwritten documents. In: Tenth international workshop on frontiers in handwriting recognition. Suvisoft
45. Louloudis G, Gatos B, Pratikakis I, Halatsis C (2008) Text line detection in handwritten documents. Pattern Recognit 41(12):3758–3772
46. Louloudis G, Gatos B, Pratikakis I, Halatsis C (2009) Text line and word segmentation of handwritten documents. Pattern Recognit 42(12):3169–3183
47. Louloudis G, Gatos B, Halatsis C (2007) Text line detection in unconstrained handwritten documents using a block-based Hough transform approach. In: Ninth international conference on document analysis and recognition (ICDAR 2007), vol 2. IEEE, pp 599–603
48. Saha S, Basu S, Nasipuri M, Basu DK (2010) A Hough transform based technique for text segmentation. J Comput 2(2):134–141
49. Abuhaiba ISI, Datta S, Holt MJJ (1995) Line extraction and stroke ordering of text pages. In: Proceedings of 3rd international conference on document analysis and recognition, vol 1. IEEE, pp 390–393
50. Sesh Kumar K, Namboodiri AM, Jawahar CV (2006) Learning segmentation of documents with complex scripts. In: Computer vision, graphics and image processing. Springer, Berlin, Heidelberg, pp 749–760
51. Kumar J, Doermann D (2010) Handwritten Arabic text line segmentation using affinity propagation. In: Proceedings of the 9th IAPR international workshop on document analysis systems, pp 135–142
52. Razak Z, Zulkiflee K, Salleh R, Yaacob MH, Tamil EM (2007) A real-time line segmentation algorithm for an offline overlapped handwritten Jawi character recognition chip. Malays J Comput Sci 20(2):171–182
53. Yin F, Liu CL (2009) Handwritten Chinese text line segmentation by clustering with distance metric learning. Pattern Recognit 42(12):3146–3157
54. Vo QN, Kim SH, Yang HJ, Lee GS (2018) Text line segmentation using a fully convolutional network in handwritten document images. IET Image Process 12(3):438–446
55. Renton G et al (2017) Handwritten text line segmentation using fully convolutional network, pp 5–9

56. Garg NK, Kaur L, Jindal MK (2010) A new method for line segmentation of handwritten Hindi text. In: ITNG2010—7th international conference on information technology: new generations, pp 392–397
57. Deshmukh MS, Patil MP, Kolhe SR (2018) A hybrid text line segmentation approach for the ancient handwritten unconstrained freestyle Modi script documents. Imaging Sci J 66(7):433–442
58. Kulkarni S, Borde P, Manza R, Yannawar P (2017) Text line segmentation of handwritten historical MODI documents. In: International conference on information, communication and computing technology (ICICCT2017), pp 1–12

Designing a Digital Shadow for Pasture Management to Mitigate the Impact of Climate Change

Ntebaleng Junia Lemphane, Rangith Baby Kuriakose, and Ben Kotze

Abstract Pasture is a said to be a natural feed to the animal. Pastures are maintained to ensure a good quality and quantity of the feed to get the best produce from the animals. Traditional ways of managing the pastures are costly and labor intensive. The use of technology in farming introduced the concept of smart farming which made pasture management easier and cost effective. Smart farming integrates advanced technological methods that include Internet of Things (IoT), big data and cloud computing. These enable tracking, monitoring, automating and complex operational analysis operations. However even with the use of these new technological techniques there are still challenges facing pasture management. Pastures are dependent on climatic conditions therefore climate change has a great impact on pasture management. With the onset of climate change, it has become a challenge for livestock farmers to predict weather conditions therefore it is impossible for decision makers to deal more effectively with the effects of climate variability. However, the digital twin concept proved to be the best technique to tackle these challenges due to its complex prediction techniques that include artificial intelligence (AI) and machine learning. This paper analyses the models used in building a digital shadow as the first step in developing a digital twin.

Keywords Artificial neural network · Climate change · Digital twin · Digital shadow · Linear regression · Smart farming

N. J. Lemphane (✉) · R. B. Kuriakose · B. Kotze
Central University Technology, Bloemfontein 9300, Free State, South Africa
e-mail: 213081768@stud.cut.ac.za

R. B. Kuriakose
e-mail: rkuriako@cut.ac.za

B. Kotze
e-mail: bkotze@cut.ac.za

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_35

1 Introduction

Pasture management is an important aspect of farm management. Pasture management is an exercise of managing healthy grass and associated plants to provide sustainable feed to the animals to ensure profitable livestock production [1]. In pasture-based farming methods, production is firmly linked to the growth and utilization of the pasture. A well-managed pasture provides notable benefits that include enhanced forage yields and reduced feed costs thereby leading to healthy livestock. Sheep farming is widely practiced in South Africa and makes a significant contribution in wool and meat production [2].

Technological developments have brought new ways to enhance management strategies in pastures for better production. This is due to the concept of smart farming. Smart farming utilizes sensors that allow remote monitoring, virtual fencing, provide tools for weather and soil information as well as inspect grazing systems [3].

Several technological methods have been put in place to properly manage pastures, such as GPS tracking, virtual fencing, and animal weighing systems [4–6]. However all these systems have the drawbacks that include linking the integrated sensor's data to the analytics propelling automation and feedback activities [7].

Digital twins offer a solution to these problems by enabling smooth integration between IoT sensors and data analytics [8]. Thus digital twin technology assures to assist livestock farmers to predict animal behaviors, monitor and prevent diseases, improve breeding operations together with management strategies to prevent livestock losses [9].

This paper discusses how to design a digital shadow as the initial step toward building a digital twin to improve management strategies in pastures to alleviate the impact of climate change.

2 Fundamentals and Related Work

Climate change is an ongoing effect that is inimical to farming and livestock production. This has resulted in emerging patterns of seasonal change, increasing temperatures and unpredictable rain [10]. This variability has caused a great harm in pasture management. However, digital twins proved to be the best solution.

A digital twin is the virtual copy of the physical object that allows bi-directional data flow between the virtual platform and the physical object [11]. Digital twins are very helpful in today's world as they can predict possible challenges before they occur. Digital shadow is the key concept in digital twin development.

Digital shadow is a predecessor to digital twins. Digital shadows have one-way data flow between the physical and digital platforms, compared to the two-way exchange in digital twins [12]. A digital shadow maps the changes in the physical objects on the digital platform, not the vice versa [13]. It assures to offer potential in utilizing predictive analytics by using complex simulation models [14].

Digital shadows are used for structuring different kinds of data collected from different data sources that can be stored, analyzed and used for various applications [15]. Digital shadows are useful in pasture management as they are used in processing farm data to improve management strategies.

This work focuses on building a digital shadow that represent unidirectional flow of data processing from the physical to the digital object. To design an effective digital shadow, powerful models are derived to allow automation and utilization of real-time data.

Digital shadow is represented by various computer-aided design model categories and engineering simulation methods [16]. Predictive modeling is the approach that uses machine learning to predict the future outcome by using historical and the current data. The model uses an algorithm that analyses both the historical and existing data to forecast the outcome. Regression and neural models are the most used predictive modeling methods [17].

Regression analysis is the method that determines the coloration among variables and use these colorations to make predictions [18]. Linear regression is the most used type of regression analysis. A model can have one independent variable or more than one independent variables [19]. A simple linear regression model can be represented with as

$$\hat{y} = \mathcal{M}_0 + \mathcal{M}_1 X + \varepsilon \tag{1}$$

where $\hat{y}$ is the predicted value of y and X is the independent variable. The terms $\mathcal{M}_0 + \mathcal{M}_1$ are the regression coefficients where $\mathcal{M}_0$ is the intercept parameter and $\mathcal{M}_1$ represents the slope. The vague error element ε accounts for the variance in the observed values and predicted values.

Artificial neural network(ANN) is a non-linear predictive method that has robust pattern classification and pattern recognition abilities to perform forecasting [20]. ANN is a data-driven self-adaptive technique thus this makes it pliable modeling mechanism for forecasting. ANN is built up by using various artificial neurons where the output of each component result to be the input of the succeeding processing component [21]. ANN structure is composed of input layer, hidden layer and the output layer. A hidden layer represents the interactivity among the neurons. ANN can be expressed as

$$Y_x = g\left(\sum_i W_{ix} M_i - \theta_x\right) \tag{2}$$

where Y_x is the output signal, g is the transfer function, W_{ix} is the link weighting, M_i is the input signal and θ_x is the bias.

3 Research Methodology

The impact of climate change on pasture-based livestock systems have been fairly given a little attention as compared to crop production despite its effects, the reason may be that managing climate change in pastures is quite complex due to various interconnecting effects [22]. A digital twin solves climate change challenge by implementing model-based predictive control to make predictions about the future [23]. Therefore, the study focuses on developing a digital shadow that predicts the climatic conditions based on four parameters, namely: temperature, rainfall, soil moisture and height of the pasture as the first step in building a digital twin.

The research methodology is split into two phases. The first phase in the design was data collection. Before the model could be developed, data to be used for the analytics was collected. The second phase of the methodology focuses on selecting models to make predictions. The model analyses the current and historical data by deploying techniques that integrate machine learning and AI algorithms [17].

3.1 Data Collection

A decade (2011–2020) historical data based on temperature and rainfall was compiled. Data was collected from world weather online databases for Bloemfontein region in South Africa. Data samples were taken for the month of June for all 10 years.

Farm data collection was also conducted. The farm is about 30.74 ha that practices mixed farming systems with integrated crop and pasture-based livestock production. This combination increases the sustainability of the farm and environmental benefits are obtained through this system, this leads to enhanced land use and rising income [24]. The study was conducted on a pasture of oats with the size of 0.64 ha.

The pasture was occupied by 45 Dohne Merino sheep. Dohne Marino is renowned as a wool breed and produces quality and weighty fleeces [25]. Data collected include pasture height, soil moisture, rainfall, temperature, border (section on rotational grazing), sheep weight and additional feed to sheep. Data was recorded for the period of three months, April–June 2021.

3.2 Selecting Appropriate Features for Developing a Predictive Model

The second phase, as described in the introduction, was to filter the data to be used and find the best suited model for development. Matrix Laboratory (MATLAB) was chosen to be the best platform to develop the model due to advantages it has

over other conventional computer languages. These include MATLAB being a high-performance language that incorporates a wide range of computing, immediate visualization and a programming environment with complex data structures [26]. The following steps were undertaken:

- Analyzing the trends in the collected data.
- Performed the comparison and overall accuracy between linear regression model and neural network model in MATLAB, to evaluate proper features to build an efficient model.
- Use the best model to predict the parameters based on the data collected on the farm.

4 Results and Analysis

4.1 Data Analysis

Data has been analyzed to realize the trend on change in temperature and the amount of rainfall over the past 10 years (2011–2020) for the month of June.

Figure 1 shows that there was an increase in temperatures from 2011 to 2013 for this month. From 2014 the temperatures started to fluctuate until 2016. From 2017 there is a slight increase in temperatures each year until 2019, and there is a slight drop in 2020. The analysis also shows that there is a significant decline in rainfall since 2011.

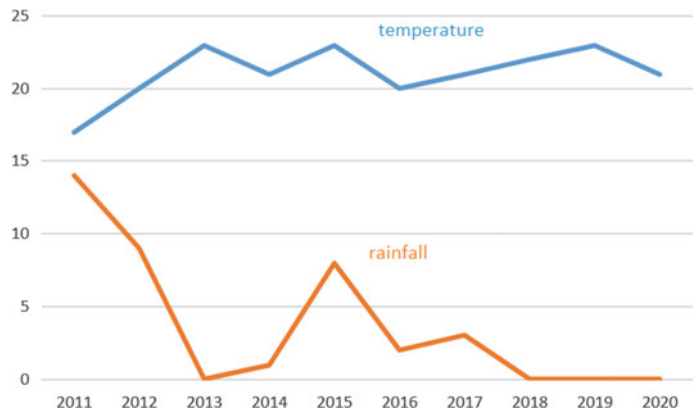


Fig. 1 Temperature and rainfall trend for the past decade

4.2 Comparing Different Predictive Models

Linear regression model was trained using the temperature statistics for the past decade (2011–2020). Data for 2021 is also used as the current data. The model was used to predict temperatures for June 4, 2021. The predictions were done for the entire day after every 3 hours. The results are shown in Table 1.

Performance metrics are important elements of error evaluation in various frameworks [27]. One of the most frequently used error measures is root mean square error (RMSE). RMSE measures the deviation between values predicted by the model and the observed values [28]. Lower RMSE indicates a better fit. MATLAB showed that linear regression model has RMSE value of 2.17 for analyzing and predicting the values. Figure 2 illustrates that the model performance does not differ that much from the actuality.

Sreehari and Ghantasala support that regression and ANN are commonly used models for climate prediction [29]. Simple linear regression (SLR) and multiple linear regression (MLR) methods are compared in predicting future climatic conditions.

Table 1 Linear regression predicted values

Time	Observed values (°C)	Predicted values (°C)
00:00	6	8
03:00	4	7
06:00	3	8
09:00	7	11
12:00	14	15
15:00	15	16
18:00	11	12
21:00	8	9

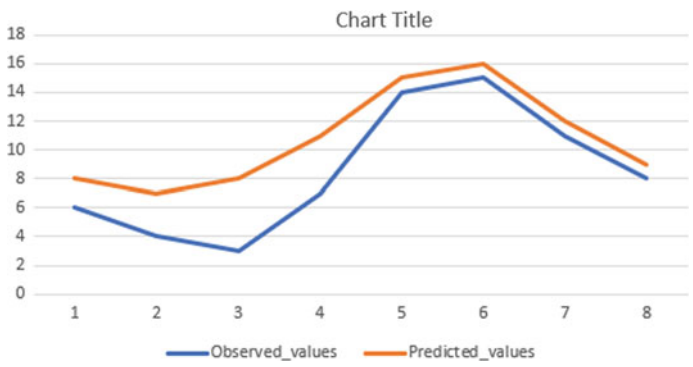


Fig. 2 Linear model performance

Table 2 ANN predicted values

Time	Observed values (°C)	Predicted values (°C)
00:00	6	5
03:00	4	5
06:00	3	5
09:00	7	7
12:00	14	13
15:00	15	15
18:00	11	11
21:00	8	8

Simple regression method proved to work better as it is most suitable in working with less complex data [29].

Similar data that was used to build the linear regression model was used build an ANN model. The results are demonstrated in Table 2.

The results in Table 2 as well as the analysis in Fig. 3 prove that ANN is more efficient than the linear regression model. Additionally, MATLAB showed ANN's RMSE value to be 1.89 which is slightly lower than of the regression model.

In comparison of ANN model and linear regression models, ANN model outperform linear regression models [30]. The proposed models were built to predict rainfall in using ANN model and MRL model. ANN produced better results than MRL due to its sophisticated processing and predictive analysis [31]. The advantage of using ANN over MRL for weather prediction is that ANN can perform operations that linear program cannot do due to its nonlinearity inherent networks that assure confidence in decision making [31].

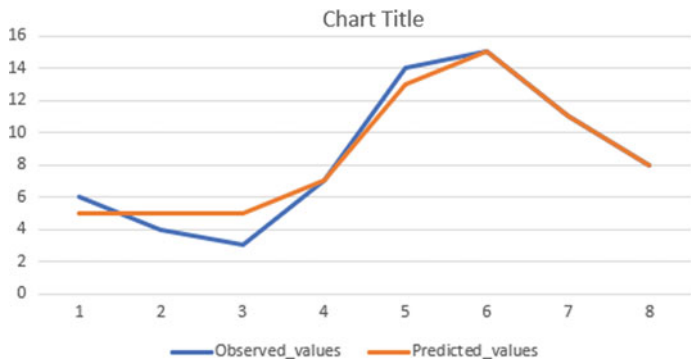


Fig. 3 ANN performance

Table 3 Predicted parameters based on farm data

Parameters	Observed values	Predicted values
Max temperature (°C)	20	19.43
Soil moisture	0	0.2
Pasture height (cm)	45	43.30
Sheep weight (kg)	37	36.42

4.3 Develop a Digital Shadow Based on Farm Data

One of the objectives of the study is to predict climate patterns so that a farmer gets prepared and make quick decisions to maintain the quality of the pasture. ANN model proved to be the better tool for prediction. Therefore, the digital shadow was developed using ANN modeling to predict temperature, soil moisture, pasture height, and sheep weight for June 28, 2021. The model was trained using the data collected from May 17, 2021. The digital shadow produced the results.

Table 3 shows that the model predicted the values which are much closer to the measured values. The correlation between the two sets of data was performed and returned a value of 0.999833. Correlation coefficient represents how closely two variables covary, the strength of the relation lies between -1 and $+1$ [32]. The value closer 1 shows the strong relationship between values.

5 Conclusion

This paper investigated ANN and linear regression models to find appropriate predictive model to build a digital shadow. Therefore, ANN proved to be best model. Nonetheless, to build a digital shadow, each output parameter is run separately because parameters have different predictors, thus this slows down the development process.

The future scenario is to use this digital shadow to develop a digital twin which will be used to predict climatic conditions. A digital twin will combat the impact of climate change in pasture management as it will assist farmers to plan accordingly.

The outstanding advantage of a digital twin is to predict unexpected scenarios, thus helping farmers to come up with quick decisions to prevent the downtime. However, this technology is dependent on Internet connection and if the connection fails, proper data analysis will be interrupted.

References

1. Grass/pasture management | grasslands beneficial management practices. http://www.nagrasslands.org/category/beneficial-management-practices/grasspasture-management/. Accessed 12 Aug 2021

2. Kom L (2016) Sheep flock structure, dynamics, management practices, and wool production under bush encroached and non-encroached areas of the Eastern Cape Province, South Africa. Euphytica 18(2):22280
3. Greenwood PL, Kardailsky I, Badgery WB, Bishop-Hurley GJ (2020) 381 smart farming for extensive grazing ruminant production systems. J Anim Sci 98(Supplement_4):139–140. https://doi.org/10.1093/jas/skaa278.257
4. Santos S, Gonçalves P (2021) Precision agriculture for crop and livestock farming—brief review, pp 1–18
5. Anderson DM (2007) Virtual fencing past, present and future. Rangel J 29(1):65–78. https://doi.org/10.1071/RJ06036
6. Dairy Australia (2014) Precision dairy technology: walkover weigh scale systems. Dairy Australia Limited. [Online]. Available: http://www.dairyaustralia.com.au/~/media/Documents/Animal management/Technologies/Animal technologies/Precision dairy technology - Walkover weigh scale systems - Case study - Mark Billing.pdf
7. What are the challenges of building a smart agriculture system? https://internetofthingsagenda.techtarget.com/answer/What-are-the-challenges-of-building-a-smart-agriculture-system. Accessed 23 Aug 2021
8. Fuller A, Fan Z, Day C, Barlow C (2020) Digital twin: enabling technologies, challenges and open research. IEEE Access 8:108952–108971. https://doi.org/10.1109/ACCESS.2020.2998358
9. Neethirajan S, Kemp B (2021) Digital twins in livestock farming. Animals 11(4). https://doi.org/10.3390/ani11041008
10. WFPC-ADAPT (2015) The impacts of climate change on food security and livelihoods in Karamoja. In: CCAFS climate change, agriculture and food security, p 52. [Online]. Available: https://reliefweb.int/sites/reliefweb.int/files/resources/wfp289992.pdf
11. Jones D, Snider C, Nassehi A, Yon J, Hicks B (2020) Characterising the digital twin: a systematic literature review. CIRP J Manuf Sci Technol 29:36–52. https://doi.org/10.1016/j.cirpj.2020.02.002
12. Sepasgozar SME (2021) Differentiating digital twin from digital shadow: elucidating a paradigm shift to expedite a smart, sustainable built environment. Buildings 11(4). https://doi.org/10.3390/buildings11040151
13. Kritzinger W, Karner M, Traar G, Henjes J, Sihn W (2018) Digital twin in manufacturing: a categorical literature review and classification. IFAC-PapersOnLine 51(11):1016–1022. https://doi.org/10.1016/j.ifacol.2018.08.474
14. Ehrhardt JM, Hoffmann CT (2020) The digital shadow: developing a universal model for the automated optimization of cyber-physical production systems based on real-time data. Procedia CIRP 93:304–310. https://doi.org/10.1016/j.procir.2020.03.069
15. Ladj A, Wang Z, Meski O, Belkadi F, Ritou M, Da Cunha C (2021) A knowledge-based digital shadow for machining industry in a digital twin perspective. J Manuf Syst 58(PB):168–179. https://doi.org/10.1016/j.jmsy.2020.07.018
16. Romero D, Wuest T, Harik R, Thoben KD (2020) Towards a cyber-physical PLM environment: the role of digital product models, intelligent products, digital twins, product avatars and digital shadows. IFAC-PapersOnLine 53(2):10911–10916. https://doi.org/10.1016/j.ifacol.2020.12.2829
17. Vaibhav Kumar ML (2018) Predictive analytics: a review of trends and techniques. Int J Comput Appl 182(1):31–37. https://doi.org/10.5120/ijca2018917434
18. Olive DJ (2017) Linear regression, pp 1–494. https://doi.org/10.1007/978-3-319-55252-1
19. Gupta BC, Guttman I, Jayalath KP (2020) Multiple linear regression analysis. In: Statistics and probability with applications for engineers and scientists using MINITAB, R and JMP, pp 693–756. https://doi.org/10.1002/9781119516651.ch16
20. Zhang G, Eddy Patuwo B, Hu MY (1998) Forecasting with artificial neural networks: the state of the art. Int J Forecast 14(1):35–62. https://doi.org/10.1016/S0169-2070(97)00044-7
21. Huang H (2013) Using artificial neural networks to establish a customer-cancellation prediction model. Prz Elektrotech 89(1b):178–180

22. Ghahramani A et al (2019) Climate change impact, adaptation, and mitigation in temperate grazing systems: a review. Sustainability 11(24):1–30. https://doi.org/10.3390/SU11247224
23. Semeraro C et al (2021) Digital twin paradigm: a systematic literature review. To cite this version: HAL Id: hal-03218786
24. Dahal S (2011) Sustainability in pasture-based livestock production system: a review. Sustainability. https://doi.org/10.13140/RG.2.2.25258.41928
25. der Merwe DAV, Brand TS, Hoffman LC (2020) Wool production in Dohne Merino, Dormer, Merino and South African mutton merino lambs. Arq Bras Psicol 50(6):881–889. https://doi.org/10.4314/sajas.v50i6.15
26. Houcque D (2005) Introduction to MATLAB for engineering students
27. Botchkarev A (2018) Performance metrics (error measures) in machine learning regression, forecasting and prognostics: properties and typology, pp 1–37. [Online]. Available: http://arxiv.org/abs/1809.03006
28. Katari MS, Tyagi S, Shasha S (2021) Statistics is easy: case studies on real scientific datasets. Synth Lect Math Stat 13(3)
29. Sreehari E, Ghantasala GSP (2019) Climate changes prediction using simple linear regression. J Comput Theor Nanosci 16(2):655–658. https://doi.org/10.1166/jctn.2019.7785
30. Hallman J (2019) A comparative study on linear regression and neural networks for estimating order quantities of powder blends
31. Soares dos Santos T, Mendes D, Rodrigues Torres R (2016) Artificial neural networks and multiple linear regression model using principal components to estimate rainfall over South America. Nonlinear Process Geophys 23(1):13–20. https://doi.org/10.5194/npg-23-13-2016
32. Mukaka MM (2012) Statistics corner: a guide to appropriate use of correlation coefficient in medical research. Malawi Med J 24(3):69–71

Parsing Expression Grammar and Packrat Parsing—A Review

Nikhil S. Mangrulkar, Kavita R. Singh, and Mukesh M. Raghuwanshi

Abstract Ford presented Parsing Expression Grammars (PEGs) as an alternative to specify rules for programming language, along with a Packrat parser, based on an idea of memoization. The idea proposed by Ford guarantees parsing of grammar written using PEGs in linear time in spite of backtracking. The primary aim of the paper is to highlight the details of PEGs followed by various challenges existing for a better understanding of the readers. From the entire overview presented, it has been observed that PEGs address the issue of undesired ambiguity in grammar for computer-oriented programming language, by not allowing any ambiguity in rules itself at the first place. However, the guarantee of linear time execution comes with a cost of large heap consumption, making it infeasible to implement for large inputs. Optimizing the resources required for memoization, may allow us to utilize the benefits offered by PEGs.

Keywords Parsing · Packrat parsing · Parsing expression grammar · PEGs · Memoization · Liner time parsing

1 Introduction

All languages that we commonly use today are originated from the idea of communicating the information as an arrangement of written signs or symbols, and which is true for both human-readable and machine-readable languages. To represent the

N. S. Mangrulkar (✉) · K. R. Singh
Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India
e-mail: nmangrulkar@ycce.edu

K. R. Singh
e-mail: ksingh@ycce.edu

M. M. Raghuwanshi
Symbiosis Institute of Technology, Symbiosis International (Deemed University), Nagpur, Maharashtra, India
e-mail: director@sitnagpur.siu.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_36

text while writing in a language, we usually use an arrangement of characters that are chosen from any standardized character set. This sequence of chosen characters makes a string in electronic form. The prime goal of any language processing application is to analyze the input strings into meaningful upper-level groupings such as words, phrases, sentences, or statements.

For deriving useful upper-level structural information from input string—also known as "syntax analysis" or "parsing", a tool called Parser is used. A grammar is commonly used for understanding and expressing a language's syntactic structure, which is a summarized depiction of the formation of a language, shown in some other language.

Every computer programming language has its own syntax which a programmer needs to follow while writing the code. Programming language syntax is a group of rules that defines the amalgamations of symbols or signs that are treated as accurately structured program fragment or program in that language. A programming language is described by the combination of its syntax and its semantics. Syntax specifies the way of writing statement, i.e., the structure of the statement, and semantics gives meaning of the statement in that programming language.

Rest of the paper is arranged as follow: Firstly, the paper presents the grammars and discuss about its related history. Later, the paper discusses about parsers and memoization technique used by Packrat parsers followed by survey of the research carried out worldwide by various researchers in consequent section of the paper. Finally, the paper discusses various issues and challenges presented in literature followed by concluding remarks.

2 History of Grammars

Chomsky, in 1950 introduced phase structure grammar, which was continuation of former work done by Alan Turing, Axel Thue, and Emil Post. In the domain of natural language representation, Chomsky suggested various simplifications and abstractions, which quickly became the standard for defining syntax of programming language.

Chomsky specified 4 types of grammars: Unrestricted grammar, Context-sensitive grammar, Context-free grammar, and Regular grammar; also known as Type-0, Type-1, Type-2, and Type-3 grammar, respectively.

All formal grammars are included in Type-0 grammars. All the languages recognizable by a Turing machine, also known as the Turing-recognizable or recursively enumerable languages, are generated precisely by this grammar.

Context-sensitive languages are generated by Type-1 grammars. The form of Type-1 grammars is depicted in Fig. 1.

Fig. 1 Type-1 grammar

$$\alpha A \beta \rightarrow \alpha \gamma \beta$$

Fig. 2 Type-2 grammar

$$A \rightarrow \alpha$$

$$A \leftarrow a\,b \,/\, a$$

Fig. 3 A parsing expression grammar

In Fig. 1, A represents a nonterminal and α β γ are strings of terminals or/and nonterminals. All the languages recognizable by a linear bounded automaton can be precisely described by Type-1 grammars.

Context-free languages are generated by Type-2 grammars. The rule for defining the grammars of Type-2 takes the form as depicted in Fig. 2.

In Fig. 2, A is a single nonterminal symbol and α is a string of symbols. All the languages recognizable by a nondeterministic pushdown automaton can be precisely described by Type-2 grammars.

Regular languages are generated by Type-3 grammars. Type-3 grammar limits its rules to a single nonterminal on the left side and a single terminal, sometimes preceded or followed by a single nonterminal on the right side of the rule. A rule is said to be Left Regular if a nonterminal precedes the terminal on right side of the rule and it is said to be Right Regular if terminal is followed by a nonterminal on right side of the rule. All the languages recognizable by a finite state automaton can be precisely described by Type-3 grammars. Furthermore, regular expressions can also be used to obtain this group of formal languages. Use of regular language is very common for defining search patterns and lexical structure of the programming languages.

3 Parsing Expression Grammars

Parsing Expression Grammars (PEGs) is an alternative that can be used for describing syntax for machine-oriented languages. "/" is a prioritized choice operator used by PEGs for specifying possible expansions for a nonterminal. The "/" operator gives multiple patterns to be examined orderly, certainly using the first successful match. A rule using PEGs can be define by using form depicted in Fig. 3.

PEG can be seen as a proper representation of a top-down parser. A memoizing parser can be used to parse all PEGs in linear time.

4 Memoization in Parsing

Memoization is a method used for optimization; mainly to increase the execution speed of computer programs by storing the results of time-consuming function calls

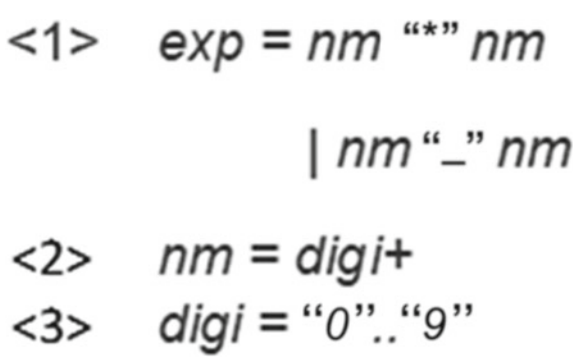


Fig. 4 Sample rules for a recursive descent parser

and returning the stored result on occurrence of the same input. To understand the concept, consider the example shown in Fig. 4.

A recursive descent parser starts matching the input "65 – 3" with the exp rule (Rule <1> in Fig. 4). It starts by comparing with the first possibility: *nm* "*" *nm.*

The sequence of one or more digits matches with the first term, *nm.* The first two symbols on the input string ("65") are successfully consumed by *nm.* Succeeding, the parser tries to match the "*" symbol, which fails as the next symbol on input string is "–". Failing to match the input symbol makes the parser to backtrack where the current possibility started. In this example, position 0.

The parser then tries to parse the input using second possibility (second line of Rule <1> in Fig. 4): *nm* "–" *nm.*

At this position, a normal parse based on recursive descent approach would apply the *nm* rule again, to match and consume the symbols at positions 0, and 1, duplicating the work that was already done while trying first possible rule. In memoization, the result of applying *nm* at position 0 is stored (memorized) on the first attempt. As the parser "remembers" that *nm* succeeds at position 0 after consuming two symbols, it can simply update the position to 2 and try to match the next symbol on input string ("–"). Finally, after successfully matching the "–" symbol, the parser applies *nm* at position 3, which consumes the last input symbol ("3") causing the successful parsing of entire exp.

A packrat parser uses concept of memoization to parse a PEG which guarantees its execution in linear time.

5 Packrat Parser

Packrat parser is a top-down approach-based recursive descent parsing with memoization. It is able to parse all CFGs in guaranteed linear timing. Packrat parsers stores all of the results evaluated which were used for backtracking (called as memoization). Memoization in packrat parsing eradicates the drawback of general top-down approach-based algorithms with backtracking, which in worst case suffers

from an exponential parsing time. The memory required for memoization is linearly proportionate to the input string size.

Thus, in this section we presented a brief overview of memoization and Packet parser which will help the reader to understand the diverse work presented by various researchers presented in forthcoming literature survey section.

6 Literature Survey

Birman et al. did the groundwork work for Packrat parsing in 1970. Birman introduced the schema called TMG recognition schema (TS) [1]. This was later polished along with Ullman, as generalized TMG recognition schema (gTS) [2]. Ford introduced Parsing Expression Grammars or PEGs in which prioritized choice operator was introduced to specify alternative expansions of a nonterminal symbol for specifying language syntax [3]. Top-down approach-based Packrat parsing was introduced by Ford, in 2002 in which concept of memoization has been implemented which guaranteed execution in linear time, despite backtracking [4, 5]. Medeiros et al. mentioned that Packrat has linear space complexity; i.e., memory required by Packrat goes on increasing linearly as length of input string increases, which makes Packrat parsers unfeasible for parsing if input data is large [6].

The work carried by different researchers can be categorized in two categories: One in which researchers have tried to enhance the capabilities of Packrat parser and second in which researchers have tried to improve or optimize the technique of memoization used by the Packrat parser.

Attempt are made by various researchers to enhance the capabilities of Packrat parser. Oikawa et al. presented an algorithm named continuation-based conversion to convert regular expressions into PEGs [7]. Becket et al. discussed that in general, Packrat parsing may be remarkably less efficient than plain recursive descent with backtracking, but memoizing selected number of recognizers can sometimes improve its performance [8]. Rats! introduced by Grimm, is a parser generator for Java programming language. Rats! supports easy modification of grammars and dodges the complexities associated with changing LL or LR grammars [9]. A tool written in Java, named Mouse, having capability of transcribing PEG into an executable parser was introduced by Redziejowski [10]. One of the snags in PEGs was inability to write PEGs equivalent to DFAs. Chida et al. presented Linear PEGs (LPEGs), a subclass of PEGs that were equivalent to DFAs [11].

Dubroy et al. developed an incremental parser by making simple modification in a standard Packrat parsers memoization strategy [12]. Incremental parsers can perform syntax analysis after each edit operation without completely reparsing the input, which is a basic requirement for any interactive code editors and IDEs. Supported by the results of their experiments, researchers claimed that their technique gave better performance for JavaScript as compared to hand-optimized, non-incremental parser [12]. Laurent et al. introduced a PEG library allowing support to left-recursion, left and right associativity and precedence rules in grammar [13].

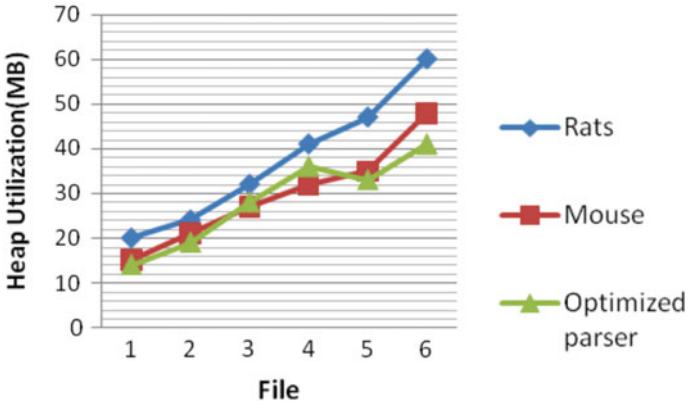


Fig. 5 Heap utilization proposed technique < heap utilization in mouse < heap utilization in rats [18]

Along with attempt to enhance the capabilities of Packrat parser, researchers have also been working on optimizing or improving the memoization technique implemented by the Packrat parsers. Moss introduced a new memoized derivative parsing algorithm which was capable of recognizing PEGs. In the proposed algorithm a concept of backtracking generation which represent backtracking decisions that occurred at some earlier point in the parse and have not yet been resolved was introduced [14]. Johnson explained how memoization can be implemented in top-down parsing and also mentioned that the memoized versions of programs derived from left-recursive grammars fail to terminate [15]. Warth et al. made modifications in memoization technique implemented by Packrat parser, to make it support left-recursive rules in grammar [16]. To be able to support left-recursive rules, modification in the memoization mechanism was done. Kuramitsu developed Nez parser, a Packrat parser based on variable buffer to minimize the heap consumption in memoization. Initially an approximate size of memoization table is selected and is changed during the run time if required. The size of table is determined as the longest backtrack required during parsing [17]. Goswami et al. implemented stack-based recursive descent parser with backtracking and with experimental results showcased improved performance over Mouse tool, a straight forward recursive descent parser [18]. Author have compared heap utilization of Rats, Mouse and their approach showing that their technique consumes the least heap among the three (Fig. 5).

7 Issues and Challenges

A parser that guarantees linear execution time along with avoiding complexity and indecisiveness of parsers based on CFGs, can be built for any PEG, but still there

are various issues and challenges related with the efficiency of Packrat parser and abilities of PEGs to represent variants of rules.

Although the ability of CFGs to express ambiguity is crucial for implementing natural/human languages, this ability literally makes it awkwardly difficult to express and to parse machine-oriented languages using CFGs [3].

Packrat parsing is a recently introduced technique which implements recursive descent parser using a top-down parsing approach. Packrat parser avoids the probable exponential time for execution of recursive descent parsing technique implementing backtracking. This is ensured by making sure that for each position in input stream, the grammar is tested only once. To achieve this, concept of memoization is implemented so that in situation of backtracking, redundant parsing can be avoided by using the memoized parsing results. Nevertheless, this benefit of guaranteed linear time execution gets nullified because of the need of large heap memory for memoization.

As memoization is a memory hungry technique, efficient use of available memory is must. Also, deciding the size of table for implementing memoization is a critical task. If the table size is too small, then we will not be able to take full advantage of what memoization can possibly provide. On other hand, if the size of table is too large, it will consume critical memory resource which will increase ratio of cache miss and thrashing. Deciding proper size of memoization table is still a challenging task.

The size of the table used for implementing memoization is also a factor for time required for searching the stored result. Time required to search an entry in the table is proportional to the number of results stored in memoization table when techniques having linear search time complexity such as binary search is implemented. Improving search time of result in memoization table can improve the performance of the parser significantly.

8 Conclusion

In this paper, authors presented an overview of how a Parsing Expression Grammars (PEGs) is an alternative that can be used for describing syntax for machine-oriented languages. PEGs solve the problem of ambiguity by not allowing any ambiguity in rules itself at the first place. A parser that guarantees linear execution time along with avoiding complexity and indecisiveness of parsers based on CFGs, can be built for any PEG. By optimizing the resources required for memoization, researchers should be able to explore the full benefits served by PEGs.

References

1. Birman A (1970) The TMG recognition schema. PhD thesis, Princeton University, Department of Electrical Engineering, Princeton, NJ

2. Birman A, Ullman JD (1973) Parsing algorithms with backtrack. Inf Control 23(1):1–34
3. Ford B (2004) Parsing expression grammars: a recognition-based syntactic foundation. In: Proceedings of the 31st ACM SIGPLAN-SIGACT symposium on principles of programming languages (POPL '04). ACM, New York, NY
4. Ford B (2002) Packrat parsing: a practical linear-time algorithm with backtracking. Master's thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA
5. Ford B (2002) Packrat parsing: simple, powerful, lazy, linear time-functional pearl. In: Proceedings of the ACM SIGPLAN international conference on functional programming, ICFP, Pittsburgh. ACM
6. Medeiros S, Ierusalimschy R (2008) A parsing machine for PEGs. In: Proceedings of the 2008 symposium on dynamic languages (DLS '08). ACM, New York, NY
7. Oikawa M, Ierusalimschy R, Moura A (2010) Converting regexes to parsing expression grammars. In: Proceedings of the 14th Brazilian symposium on programming languages, SBLP, vol 10
8. Becket R, Somogyi Z (2008) DCGs + memoing = packrat parsing but is it worth it? In: 10th international symposium, PADL 2008, San Francisco, CA, 7–8 Jan 2008
9. Grimm R (2004) Practical packrat parsing. New York University technical report, Dept. of Computer Science, TR2004-854
10. Redziejowski RR (2009) Mouse: from parsing expressions to a practical parser. In: Concurrency specification and programming workshop
11. Chida N, Kuramitsu K (2017) Linear parsing expression grammars. Preprint version, Sept 2017
12. Dubroy P, Warth A (2017) Incremental packrat parsing. In: Proceedings of 2017 ACM SIGPLAN international conference on software language engineering (SLE'17). ACM, New York, NY
13. Laurent N, Mens K (2016) Parsing expression grammars made practical. Preprint version, Sept 2016
14. Moss A (2014) Derivatives of parsing expression grammars. arXiv preprint arXiv:1405.4841
15. Johnson M (1995) Memoization in top-down parsing. Comput Linguist 21
16. Warth A, Douglass JR, Millstein TD (2008) Packrat parsers can support left recursion. In: PEPM 8
17. Kuramitsu K (2015) Packrat parsing with elastic sliding window. J Inf Process 23(4)
18. Goswami MM, Raghuwanshi MM, Malik L (2016) Performance improvement of stack based recursive-descent parser for parsing expression grammar. Int J Latest Trends Eng Technol 6(3)

Precision Testing of a Single Depth Camera System for Skeletal Joint Detection and Tracking

Ketan Anand, Kathiresan Senthil, Nikhil George, and Viswanath Talasila

Abstract Implementing a marker-less, camera-based system, to capture skeletal joint motion in a timely manner that is robust to variables such as lighting conditions, occlusion and changes in the subject's distance from the camera remains a challenge. This paper aims to test the feasibility of such a system using the Intel RealSense D435i depth camera to capture joint coordinates in the camera reference frame. The subject was placed in various lighting conditions to analyze the effect of glare on joint coordinate precision, a decimated filter and a spatial filter were applied to test which of these filters reduced noise in the image signal most effectively. To examine which setup is the most effective in measuring coordinates of the neck, elbows, wrists, knees and ankles precisely, variations were made in the subject's distance from the camera, position in the field of view and ambient lighting conditions. The empirical cumulative distribution function (ECDF) plots of the true distance obtained from the data collected were used to systematically eliminate undesirable recording conditions. The coordinates measured will be used to calculate gait parameters.

Keywords Biomechanics · Gait analysis · Joint modeling · Medical imaging · Motion capture · Data analysis · Depth camera

1 Introduction

The emerging domain of tele-rehabilitation aims to eliminate bulky equipment traditionally used in the rehabilitation of those with motor impairment. The first stage of our proposed setup of a compact, simplified module for tele-rehabilitation involves a depth camera which detects the joint coordinates of the subject. This paper focuses

K. Anand (✉) · K. Senthil
Department of Electronics and Instrumentation Engineering, M. S. Ramaiah Institute of Technology, Bengaluru, India
e-mail: anand_ketan@outlook.com

N. George · V. Talasila
Department of Electronics and Telecommunication Engineering, Center for Imaging Technologies, M. S. Ramaiah Institute of Technology, Bengaluru, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_37

on testing the feasibility of using the Intel RealSense D435i depth camera along with the *cubemos* Skeleton Tracking SDK to obtain the 3-D coordinates of the subject's joints. Joint detection and tracking are vital for measuring gait parameters. The system proposed in this paper will subsequently be used to analyze the gait of children of the age of 6–12 years, affected with cerebral palsy. A marker-less setup to track skeletal joints is especially beneficial in the process of rehabilitating those with sensorimotor impairment such as cerebral palsy. This system is easily integrable with inertial measurement units (IMUs) to develop a comprehensive gait analysis tool. Gait analysis involves measurement of joint angles, as well as joint velocity and acceleration in 3-D [1]. These gait measurements are used as an important diagnostic tool to decide the course of rehabilitation treatment in those with movement disorders.

Precision of the joint coordinates that are being measured is of importance since the data obtained is used in subsequent stages of the system proposed by us, to calculate diagnostic parameters such as gait cycles and base of support. Precision is an indicator of the reproducibility of data collected, over a given range. One metric used to determine the goodness of precision is the standard deviation, where a smaller standard deviation indicates a higher degree of precision.

The factors that were evaluated in testing the robustness of our depth camera-based system were: placing the subject in varied lighting, evaluating the performance of *cubemos* while the subject is at varied distances from the depth camera and testing various signal filters to reduce the standard deviation and improve the robustness of the system [2].

A thorough investigation of the precision of the Intel RealSense D3435i depth camera along with the *cubemos* Skeleton Tracking SDK was done by first performing camera calibration (Sect. 2.1) in indoor conditions. Further, carrying out an evaluation of the factors that influence the precision of joint coordinates (Sect. 3) helped eliminate unfavorable conditions that would affect the robustness of our gait analysis system. The quantitative results for the measured neck, right wrist, and right ankle coordinates are presented there after (Sect. 4), followed by plotting the data in the form of an Empirical Cumulative Distribution Function (ECDF) to visualize the range covered by the measured data in various scenarios (Sect. 5).

2 Methodology

2.1 Camera Calibration

The process of calibrating the depth camera ensures that lens distortion does not worsen the precision of our setup, hence the reproducibility of the coordinates is maintained.

It was also noted that points in the corner of the field of view had negligible effect on the standard deviation of any of the coordinates. Since precision of coordinates

was maintained in all regions of the field of view, calibration with the extrinsic parameters was not performed and hence, manipulating the camera coordinates with the intrinsic calibration parameters was sufficient.

Intrinsic Parameters. The intrinsic parameters of the camera are a transformation of the form $R^3 \rightarrow R^2$. This transformation maps 3-D camera coordinates to 2-D pixel coordinates [3]. The intrinsic parameter model of the Intel RealSense D435i camera is [4]:

$$\text{Proj}(x, y, z) = F \cdot D_{\text{Model}}\left(\frac{x}{z} + \frac{y}{z}\right) + P \tag{1}$$

where $D_{\text{Model:}}\ R^2 \rightarrow R^2$ is the lens distortion function, $P = \left(p_x, p_y\right)$ is the principal point indicating pixel offset from the left edge, and $F = \left(f_x, f_y\right)$ is the focal length in multiples of pixel size.

The intrinsic parameter model given above holds true for unfiltered data only. On application of signal filters such as the decimated filter or spatial edge-preserving filter, the intrinsic parameters are modified [5]. The description and controls of these filters are listed under Sect. 3.3.

Deprojection Model. The deprojection model uses the depth coordinate along with the pixel coordinates to reconstruct any point in focus, in camera coordinates. Obtaining these camera coordinates gives us the flexibility to estimate gait parameters in various biomechanical frames of reference. The deprojection model for the Intel RealSense D435i is given as [6]:

$$\text{Deproj}(i, j, d) = \left(d \cdot U_{\text{Model}}\left(\frac{(i, j) - P}{F}\right), d\right) \tag{2}$$

where $U_{\text{Model:}}\ R^2 \rightarrow R^2$ is the undo lens distortion function.

3 Evaluation

3.1 Effect of Lighting Conditions

Stereo cameras such as the Intel RealSense D435i utilize two infrared cameras and an onboard embedded D4 ASIC [7] to stream live depth data. A disparity map is generated where objects closer to the camera will have a greater horizontal shift between the right and left camera stream, compared to objects that are farther away. This shift along the epipolar line is known as disparity. The disparity is mapped to a particular depth value by a method known as *correspondence*:

$$\text{Depth} = \frac{\text{Baseline} \times \text{Focal Length}}{\text{Disparity}} \tag{3}$$

where Baseline indicates the midpoint between the left and right camera lens and Disparity is the difference in pixel value of the feature under consideration.

Improper lighting conditions that are overexposed, causes speckles in the IR pattern that the stereo cameras rely on to compute the depth coordinate. Speckles decrease depth image density, which results in incorrect correspondence. Further, if the loss of depth image density is at joints that *cubemos* is trying to detect, the *cubemos* model fails to recognize features of interest, resulting in inaccurate coordinate values [8].

3.2 Effect of Distance of the Subject from the Camera

Disparity is inversely proportional to distance of the subject from the camera. At distances with a depth component beyond 7 m the disparity becomes negligible and insensitive to changes in depth [9]. From the quantitative results presented in Sect. 4, it is observed that a subject positioned at approximately 3 m from the camera is best suited for precise depth measurement. Beyond 7 m, the negligible depth value results in failure of correspondence.

Subjects that are too close to the camera (<2 m) and not visible from head-to-toe present an equally challenging problem to our setup. When *cubemos* fails to detect a certain joint, the coordinates logged are sparse in nature.

3.3 Effect of Various Signal Filters

Image signal filters are used to reduce noise in the image data. The two filters implemented in testing our single depth camera system were the Decimated filter and the Spatial Edge-Preserving filter.

Decimated Filter. The decimated filter reduces the complexity of the depth scene by reducing the sampling frequency by an integer value. The linear scaling factor provided by the Intel RealSense post processing toolkit runs between kernel sizes of 2×2 to 8×8. Scaling the image stream with a 3×3 kernel worked best for skeleton tracking as presented in Sect. 4.

Spatial Edge-Preserving Filter. The spatial edge-preserving filter exploits the fact that an RGB image is a 2-D manifold in a 5-D space. The spatial filter defines a domain transform from $R^2 \rightarrow R^5$ [10]. A 5×5 kernel is used in this transformed space to perform smoothing and edge-preserving operations. The controls of the spatial edge-preserving filter include a filter magnitude value which indicates the number of filter iterations, the smooth alpha value defines the weight of the current

Table 1 Filter controls of the spatial edge-preserving filter

Parameter	Range	Value
Filter magnitude	[1–5]	5
Smooth alpha	[0.25–1.0]	1
Smooth delta	[1–50]	50

pixel and the smooth delta value defines the depth gradient below which smoothing will be performed. The controls of the spatial filter were set as in Table 1.

4 Quantitative Results

A thorough evaluation of the factors that influence the quality of the joint coordinates helped formulate a method of data collection that had the least standard deviation. The tabular columns (Tables 2, 3 and 4) contain statistics of the neck, right wrist and right ankle coordinates. The joint coordinates that are being measured are in the format (x, y, z, true_distance) where the true_distance of the subject from the camera origin is a preliminary indicator of the combined standard deviation of the x, y, z coordinates. Quantitative analysis was performed for the coordinates of the neck, elbows, wrists, knee and ankle, and a subset of the data is presented in Tables 2, 3 and 4.

5 Visualizing the Results

Depth Maps. The depth maps of the subject, as seen by the camera create a color coded image of pixels at the same distance. These depth maps give us useful insight into the effect of distance and glare on camera perception. The raw normalized depth map, with the joints in focus help log data more efficiently. The true_distance calculated from the (x, y, z) joint coordinates are plotted as an Empirical Cumulative Distribution Function (ECDF). The empirical distribution function is computed as:

$$F_n(t) = \frac{1}{n} \sum_{j=1}^{n} I_{\{Z_j \leq t\}} \tag{4}$$

where I is the indicator function which has a value equal to 1 if the distance for which frequency is being computed is found in the data and 0 otherwise.

Visualizing the ECDF Plots of Our Data. The ECDF plots in Figs. 1 and 2 represents the distribution of the right ankle coordinate data while the subject was at distance of 3 m and 7 m from the camera, respectively. The ECDF plot makes visualizing the

Table 2 Statistics of neck coordinates

	Actual distance from the camera = 3 m		Actual distance from the camera = 7 m		Background with glare		Evenly lit, opaque background	
	Decimated (m)	Spatial (m)	Decimated (m)	Spatial (m)	Decimated (m)	Spatial (m)	Decimated (m)	Spatial (m)
Mean	3.143	2.928	7.19	6.5	2.907	3.107	3.177	2.807
Standard deviation	0.025	0.016	0.088	0.077	0.047	0.033	0.031	0.024
Range	2.834–3.276	2.86–3.001	6.857–7.522	6.166–6.674	2.540–2.966	2.95–3.338	2.305–3.398	2.642–2.966

Table 3 Statistics of right wrist coordinates

	Actual distance from the camera = 3 m		Actual distance from the camera = 7 m		Background with glare		Evenly lit, opaque background	
	Decimated (m)	Spatial (m)	Decimated (m)	Spatial (m)	Decimated (m)	Spatial (m)	Decimated (m)	Spatial (m)
Mean	3.33	3.072	6.961	6.914	3.009	3.196	3.382	3.009
Standard deviation	0.045	0.019	0.059	0.063	0.056	0.047	0.027	0.021
Range	3.102–3.46	3.010–3.144	6.735–7.16	6.582–7.208	2.849–3.218	3.028–3.414	3.172–3.55	2.849–3.218

Table 4 Statistics of right ankle coordinates

	Actual distance from the camera = 3 m		Actual distance from the camera = 7 m		Background with glare		Evenly lit, opaque background	
	Decimated (m)	Spatial (m)	Decimated (m)	Spatial (m)	Decimated (m)	Spatial (m)	Decimated (m)	Spatial (m)
Mean	3.24	3.067	7.019	6.884	2.932	3.173	3.305	2.828
Standard deviation	0.043	0.013	0.071	0.068	0.045	0.036	0.015	0.028
Range	3.032–4.035	3.027–3.156	6.81–7.307	6.595–7.131	2.724–3.883	3.081–3.317	3.134–3.527	2.583–2.909

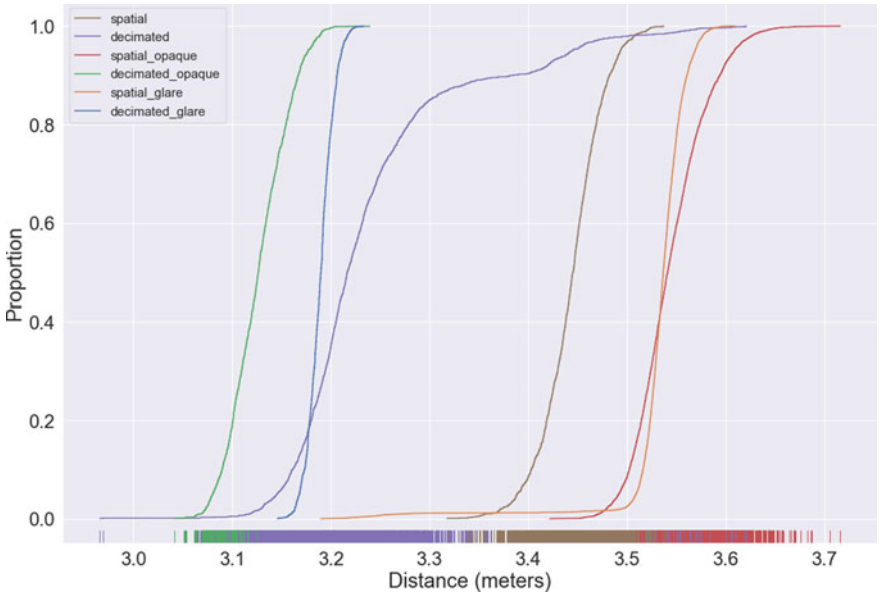


Fig. 1 ECDF plots of the right ankle coordinates logged at a distance of 3 m from the camera

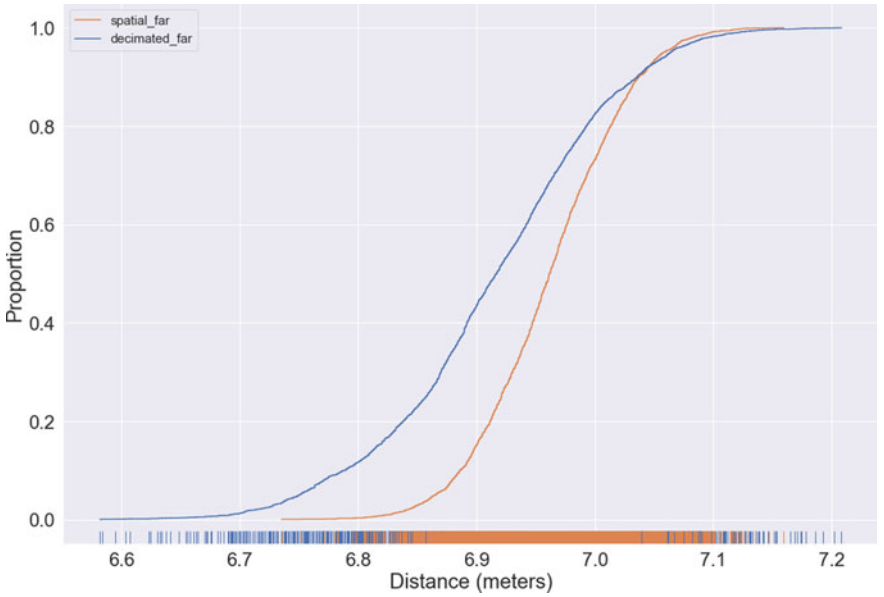


Fig. 2 ECDF plots of the right ankle coordinates logged at a distance of 7 m from the camera

range, mean and standard deviation more intuitive and making inferences simpler. The inferences drawn help us select the appropriate conditions to implement our setup in. The ECDF plots for the neck, elbows, wrists, knees and ankles aid in making sound inferences regarding appropriate recording conditions, and a subset of these plots are presented.

6 Conclusion

Rigorous analysis of the factors that could affect the performance of a single depth camera system for skeletal joint detection and tracking was performed. This yielded a set of conditions for joint coordinate measurement under which joint tracking was the most conducive. The ideal conditions for implementing our proposed setup are: no glare falling on the camera lens and the subject standing about 3 m from the depth camera, while using a spatial edge-preserving filter to post process the raw depth image captured. The standard deviation of coordinates measured in favorable, controlled conditions was under 2 cm. A high degree of precision ensures that the noise in the coordinate data is not amplified when the joint coordinates are used to calculate gait parameters, in our proposed tele-rehabilitation module. As observed from the ECDF plots and depth maps that subjects placed at a distance greater than 7 m from the camera and with overexposed lighting conditions, the precision decreases drastically. A controlled environment that is obstacle free, with the subject clearly visible in the field of view of the camera provides the most suitable conditions to perform tele-rehabilitation. Precision testing of the Intel RealSense D435i depth camera with the *cubemos* Skeleton Tracking SDK completes the preliminary stage of our rehabilitation setup, using which data can be measured and calculation of vital gait parameters can be made [11]. In future work, we aim toward integrating this single depth camera system with IMUs to create a compact system that is capable of measuring data of a variety of gait parameters [12]. Further stages of the setup we are proposing would sense and model gait using this depth camera plus IMU-based system, which would provide crucial metrics required for physical rehabilitation of gait [13].

Acknowledgements This work was supported by a BIRAC grant Ref. No. BT/AIR094S/PACE-19/19.

References

1. Raghavendra P, Sachin M, Srinivas PS, Talasila V (2017) Design and development of a real-time, low-cost IMU based human motion capture system. In: Vishwakarma H, Akashe S (eds) Computing and network sustainability. Lecture notes in networks and systems, vol 12. Springer, Singapore. https://doi.org/10.1007/978-981-10-3935-5_17

2. Chin LC, Basah SN, Yaacob S, Din MY, Juan YE (2015) Accuracy and reliability of optimum distance for high performance Kinect sensor. In: 2015 2nd international conference on biomedical engineering (ICoBE), pp 1–7. https://doi.org/10.1109/ICoBE.2015.7235927
3. Herrera CD, Kannala J, Heikkilä J (2012) Joint depth and color camera calibration with distortion correction. IEEE Trans Pattern Anal Mach Intell 34(10):2058–2064. https://doi.org/10.1109/TPAMI.2012.125
4. Intel RealSense documentation. https://dev.intelrealsense.com/docs/projection-texture-mapping-and-occlusion-with-intel-realsense-depth-cameras
5. Schreven S, Beek PJ, Smeets JBJ (2015) Optimising filtering parameters for a 3D motion analysis system. J Electromyogr Kinesiol 25(5):808–814
6. Balakrishnan G, Dalca A, Zhao A, Guttag J, Durand F, Freeman W (2019) Visual deprojection: probabilistic recovery of collapsed dimensions. In: 2019 IEEE/CVF international conference on computer vision (ICCV), pp 171–180. https://doi.org/10.1109/ICCV.2019.00026
7. Intel RealSense documentation. https://dev.intelrealsense.com/docs/post-processing-filters
8. Ling J, Tian L, Li C (2016) 3D human activity recognition using skeletal data from RGBD sensors. In: Bebis G et al (eds) Advances in visual computing. ISVC 2016. Lecture notes in computer science, vol 10073. Springer, Cham. https://doi.org/10.1007/978-3-319-50832-0_14
9. Høilund C, Moeslund TB, Madsen CB, Trivedi MM (2010) Improving stereo camera depth measurements and benefiting from intermediate results. In: 2010 IEEE intelligent vehicles symposium, pp 935–940. https://doi.org/10.1109/IVS.2010.5547978
10. Gastal E, Oliveira M (2011) Domain transform for edge-aware image and video processing. ACM Trans Graph 30:69. https://doi.org/10.1145/2010324.1964964
11. Samprita S, Koshy AS, Megharjun VN, Talasila V (2020) LSTM-based analysis of a hip-hop movement. In: 2020 6th international conference on control, automation and robotics (ICCAR), pp 519–524. https://doi.org/10.1109/ICCAR49639.2020.9108052
12. Vidyarani KR, Talasila V, Megharjun N, Supriya M, Ravi Prasad KJ, Prashanth GR (2021) An inertial sensing mechanism for measuring gait parameters and gait energy expenditure. Biomed Signal Process Control 70:103056. ISSN 1746-809. https://doi.org/10.1016/j.bspc.2021.103056
13. Megharjun VN, Talasila V (2018) A Kalman filter based full body gait measurement system. In: 2018 3rd international conference on circuits, control, communication and computing (I4C), pp 1–5. https://doi.org/10.1109/CIMCA.2018.8739597

Generation of User Interfaces and Code from User Stories

Samia Nasiri, Yassine Rhazali, Amina Adadi, and Mohamed Lahmer

Abstract Coding is time-consuming and confusing more than any other single process in a software development project. This research addresses the above issue and proposes the transition from a computation independent model (CIM) to a code whose aim is to generate cross-platform mobile and Web applications with the CRUD functionalities (create, read, update, and delete). The implementation of the model-driven architecture (MDA) approach resides in the generation of cross-platform applications, through model-to-text transformations with ACCELEO. Our automation approach is based on the agile requirement to generate an Ecore file representing the platform-independent model (PIM). Subsequently, the source code and interfaces are automatically built. The CRUD application is generated for each class of an Ecore file. These output applications are carried out using the Flutter framework. In our development process, the automation of transformations from one step to another leads to productivity gains and a reduction in overall costs.

Keywords User stories · MDA · ACCELEO · Cross-platform development · Flutter

1 Introduction

Today, mobile applications, as well as Web applications, are used everywhere in our daily life. These two types of applications are very different. Not only are there differences for the useless, but they are also developed and deployed differently. The diversity of mobile operating systems is leading companies to make more efforts and offer more services to their customers. This has pushed developers to develop cross-platform mobile applications, as each operating system requires a programming language, and development tools. The concept of cross-platform application development lets the developer write the code once and then use it on any platform [1].

S. Nasiri (✉) · Y. Rhazali · A. Adadi · M. Lahmer
Moulay Ismail University, Meknes, Morocco
e-mail: nasiri.samia@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_38

In a software life cycle, the design phase is a very critical stage in which models are established; like class diagram, use case, and others UML diagrams. These models help a lot to understand the functionalities of the information system which are complex. The design phase does not only model the data but also the creation of interfaces. Model-driven engineering focuses on placing models at the core of the development process to enable them to be the principal elements by which applications are driven. In this context, our approach presents the transition from a CIM model to a code whose aim is to generate cross-platform applications mobile and Web applications with the CRUD functionalities, based on an Ecore file automatically extracted from the agile requirements named user stories [2]. This Ecore file equivalent to a class diagram. We used the cross-platform concept by choosing the Dart language and Flutter framework to accomplish the development phase. To automate the generated code, the M2T transformation is used through ACCELEO which relies on the Ecore file. In our previous work [3], our objective was to capture the design elements that make up a UML class diagram. We started by defining several rules to detect, firstly the classes and associations, and secondly, attributes and operations classes. Our tool generates UML class diagram as a "PNG" image using a PlantUML, and also an Ecore file. This file represents PIM model in MDA approach. The transition from CIM to PIM was carried out, based on requirements expressed in user stories.

In this paper, we present a transformation from PIM to code. Our objective is to reduce the development cost of mobile and Web applications. This paper is structured as follows: In the next section, we review related work. In Sect. 3, we describe our proposal. Section 4 outlines the results, and Sect. 5 presents the evaluation. Finally, we conclude in Sect. 6.

2 Related Work

Most of researchers in field of generation user interfaces and code are focused on user requirement which are presented in use case model or other models. Plus, the rule-based transformations are defined for all model-to-model transitions and model-to-code transformations. But they did not generate user interfaces and code of the software from agile requirements. We shall now talk about some of the more relevant proposals.

In [4], the authors applied the UWE and ATL transformation language in order to generate JSP pages and JavaBeans. In [5], the approach integrates Struts2, Spring IoC, and Hibernate DAO frameworks, and its objective is to produce the N-tier PSM model. The transformation rules are determined using the ATL language. In [6], based on the MDA approach, the authors propose an approach named Model2Roo that allows to develop high-quality Web applications. A script Spring Roo has been generated illustrating the definition of a persistence layer, Web controllers, and integration tests. In [7], the authors applied the MDA approach to Web application engineering. Their approach allows to generate an XML file that is extendable to other platforms

such as Dot NET and PHP. The class diagram presents the source model. In [8], the authors present a UML diagram for modeling AJAX user interfaces based on the MDA approach. They realized an AJAX cartridge by adopting ANDROMDA. Their approach was able to generate an entire Web application based on AJAX with automatic back-end integration. The rules of transformation are implemented to achieve their goal. The generation file is an XML file that depicts the struts file. In [9], the authors exploit the UWE to provide rules and algorithms to produce code automatically from various models. They develop a code generator tool named CODEGER-UWE. In [10], a transformation of the PIM applied by enterprise distributed object into PSM is performed for diverse service platforms. A PSM model is generated. The transformation rules are defined by the ATL language. This approach does not generate the PSM model code. In the work [11], the objective is to create a metamodel describing user interfaces according to the MDE approach. The creation of the metamodel is taken from the metamodel of the graphical user interface components of the Java programming language. Their metamodel is focused on the user interface components such as the window and the frame. In [12], the code is generated by using ANDROMDA source code generator supported by MDA. Their work is done using model-based development. In [13], the authors propose a transformation from UML state machine diagram to a formal workflow language YAWL.

Our approach does not purpose to create a metamodel that represents the component of the user interface such as the majority of research, or define the transformation rules using ATL language, but we have focused on the generation of code and user interfaces using ACCELEO by developing several template files ".mtl" for each Ecore class. These files are part of the Flutter application. The Flutter technology has shown its efficiency since it allows creating a single code base for multiplatform, which reduces the time of development. Hence, the main contributions of our approach are

- The generation of the code and interfaces using ACCELEO generator.
- The generated code is CRUD applications with Flutter framework corresponding to cross-platform development.

3 Approach

Model-driven architecture (MDA) is as an architectural framework for software development [14, 15]. MDA specifies three default models that can be defined more precisely as layers of abstraction, since each of these three layers can be used to build a set of models, each model corresponding to a more specific view of the system. Indeed, MDA includes three types of transformation: computer independent model (CIM) to platform-independent model (PIM) and PIM to platform specific model (PSM) and PSM to code.

Our approach presents the transition from a CIM model to a code whose aim is to generate cross-platform applications mobile and Web applications with the CRUD functionalities. In our approach, CIM is presented by agile requirements, PIM is

presented by the Ecore file, and code is carried out by the M2T transformation with ACCELEO. This latter provides cross-platform applications Web and mobile.

A user story is a very high-level definition of a requirement, which includes just sufficient information; it is the best way to describe a requirement. Agile project teams use one of these methods: Scrum, XP, Kanban, etc. In [3], we were able to generate a class diagram from a set of user stories. This diagram is generated as a "PNG" image file using PlantUML, and as an Ecore file. The automation was accomplished using NLP tool called Stanford core NLP, Python language, and also WordNet. After the generation of an Ecore file which represents the class diagram, we proceeded to our proposal transformation. To do this, we have used ACCELEO to generate CRUD applications through "Dart" language and Flutter framework. Figure 1 shows the architecture of the proposal approach.

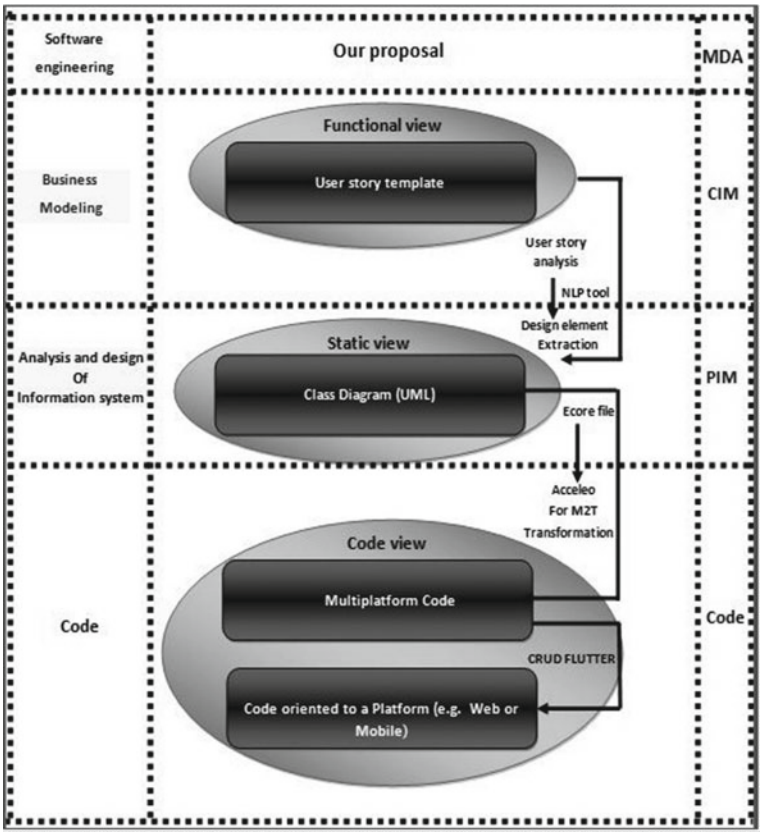


Fig. 1 Architecture of the proposal approach

3.1 Ecore Metamodel

EMF affords tools and runtime support to generate a set of Java classes for the model, by the side of a set of adapter classes that enable viewing and command-based editing of the model and a basic editor. Among the models of this framework, we quote Ecore model. This model is essentially a subset of UML class diagram.

The Ecore metamodel is a robust tool for model-driven architecture (MDA) design, which can be utilized as a basis for software development. As a general rule, we would describe the objects (of EClass type) in our area of application, their attributes, and relationships. We would also define the specific operations that are part of these objects with the help of the model element EOperation, the superclass of EAttributes and EReferences is EStructuralFeature.

3.2 Code and User Interfaces Generation

Model-driven architecture and models form the basis of a set of software development approaches called model-driven development (MDD). Models are the fundamental elements of MDA, which aims to transform platform-independent models into platform-dependent models, and ultimately into executable source code. Hence, model transformation is actually a major key in the MDA process [16, 17]. The template approach consists of defining templates which are template models or target models with parameters, these parameters will be substituted by the information contained in the source model. This approach is used in the generation of source code from templates. ACCELEO is an open template-based approach from the Eclipse Foundation. It allows the application of the model-driven architecture (MDA) approach for generating source code from EMF-based models. ACCELEO is an implementation of the MOF models to text standard of the object management group (OMG) for model-to-text transformations [18, 19].

In our case, a code source is generated using ACCELEO based on Ecore file, we noted that the Ecore file used in EMF eclipse is not created manually but it is the output of our approach in [3]. We have done a model-to-text (M2T) transformation from within the MTL template files.

A multi-layered software architecture consists of several layers, each of which corresponds to a service. Since each layer is distinct, it is easier and more efficient to make changes to each layer than to tackle the whole architecture. We have developed several MTL files, one file corresponds to the presentation layer, and the others files are related to the CRUD functionalities and form the DAL layer. Figure 2 shows the layers of the approved architecture.

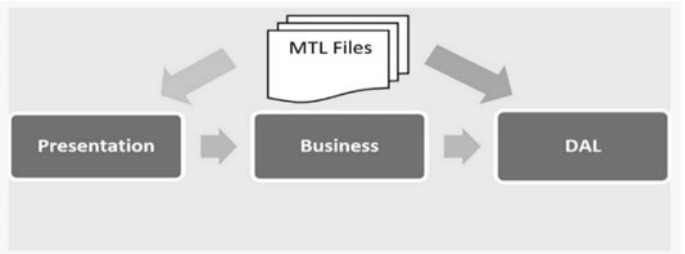


Fig. 2 Layers of the approved architecture

3.3 Flutter the Cross-Platform Development

The concept of cross-platform mobile application development has expanded considerably [20, 21]. It lets the developer write the code once and then use it on any platform Android, iOS, or Windows.

Flutter framework is belongs the cross-platform development tools that build natively compiled apps for mobile, Web, desktop, and embedded devices from a single code base. This framework is developed by Google and uses Dart programming language which is an object-oriented programming language.

4 Results

In this section, we show a case study of inline course management [3]: videos, quizzes, and others, in order to demonstrate our method of transforming PIM level to code. The Ecore file used in our evaluation corresponds to the case study that is generated from a set of user stories in our previous work [3].

To generate the CRUD application corresponding to the Ecore classes, we used ACCELEO, in the same way for the user interfaces. The CRUD actions are mostly performed in all systems. The figures displayed to illustrate a part of interfaces and code generated from an Ecore class named "event." The generated code is relevant to cross-platform Web and mobile applications. We have used Cloud Firestore database of firebase platform for storing data.

Figure 3 shows a list of events stored in firestore-cloud and depicts actions like add or update or delete a selected event in the list. Figures 4 and 6 show an interface to add a new event. Figures 5 and 7 show the generated interface to update a selected event. To this end, we have developed several MTL files using ACCELEO. The first file is part of the presentation layer which contains the components of the user interface, and each of them corresponds to an attribute of a class as shown in Figs. 4 and 6. The other files are related to the "create," "read," "update," and "delete" (CRUD) actions representing DAL layer. We note that CRUD functionalities have been realized in both mobile and Web.

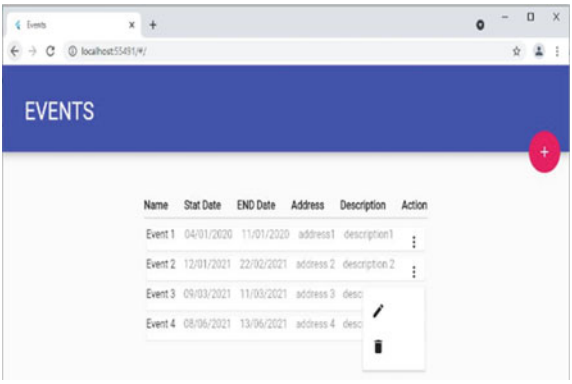


Fig. 3 Event list with actions

Fig. 4 Generated web interface to add a new event

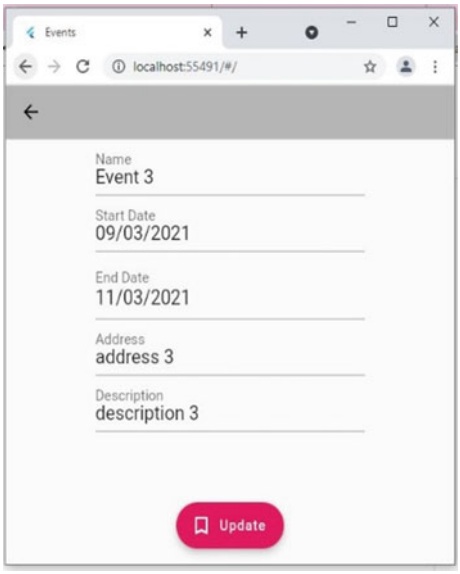


Fig. 5 Generated web interface to update an event

Figure 8 shows an ACCELEO project structure which contains the various templates to implement automatic code generation by ACCELEO. The ACCELEO project is structured as follows:

- The DAL folder containing the generated files presenting CRUD actions. Figure 9 shows a part of generated files of DAL layer.
- The presentation folder presenting the different generated files corresponding to classes of the Ecore model. Figure 10 shows a part of generated files of the presentation layer.

5 Evaluation

In this paper, we have proposed an approach for code generation using M2T transformation. This transformation has been performed using the ACCELEO generator.

Different templates have been implemented to automatically generate the code. The percentage of generated CRUD applications reached 100% of the total code requested. In this case study, we obtained a very high rate of generation of the required code.

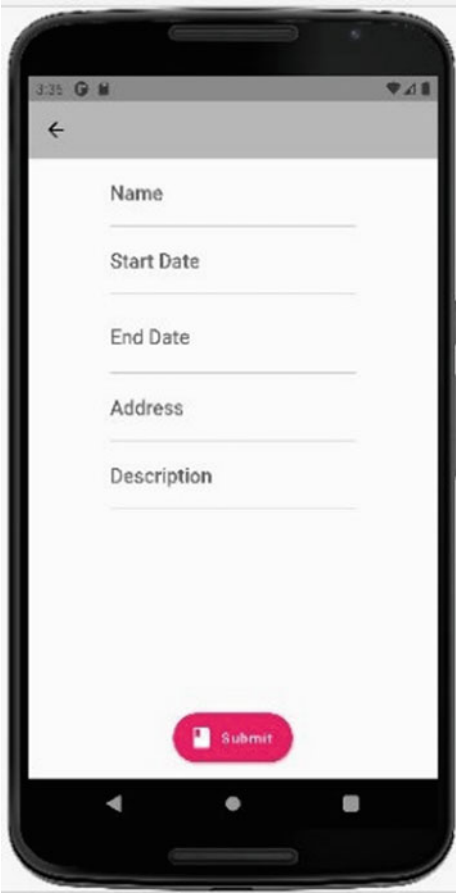


Fig. 6 Generated mobile user interface to add a new event

6 Conclusion

This paper proposed an approach to transform PIM into code and generate cross-platform Web and mobile applications with CRUD functionalities. To perform the code generation operation, we used ACCELEO as the code generation language. To accomplish this automation, we started by developing the diverse templates corresponding to each meta-class constituting the PIM model previously generated. After implementing the different templates, we validated the automatic code generation through a case study. Our approach facilitates the analytical tasks of analysts and developers to create software quickly. This is achieved thanks to the advantages of the cross-platform concept realized by the Flutter framework. As a perspective, we will focus on the business layer to generate business services based on user stories, and we will also add other functionalities related to code generation from the DAL layer by integrating referential integrity.

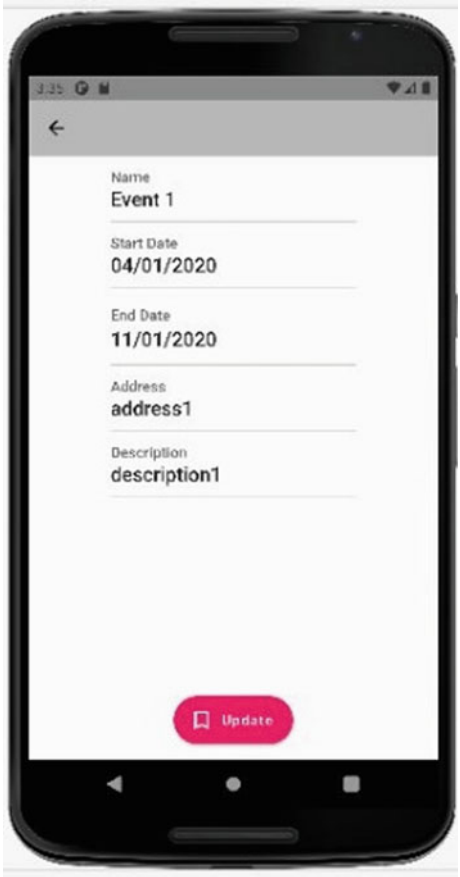


Fig. 7 Generated mobile user interface to update an event

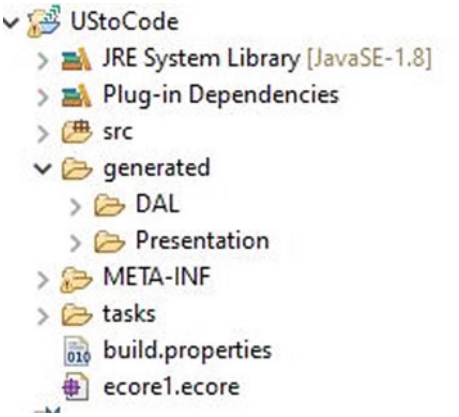


Fig. 8 Project tree of generated code

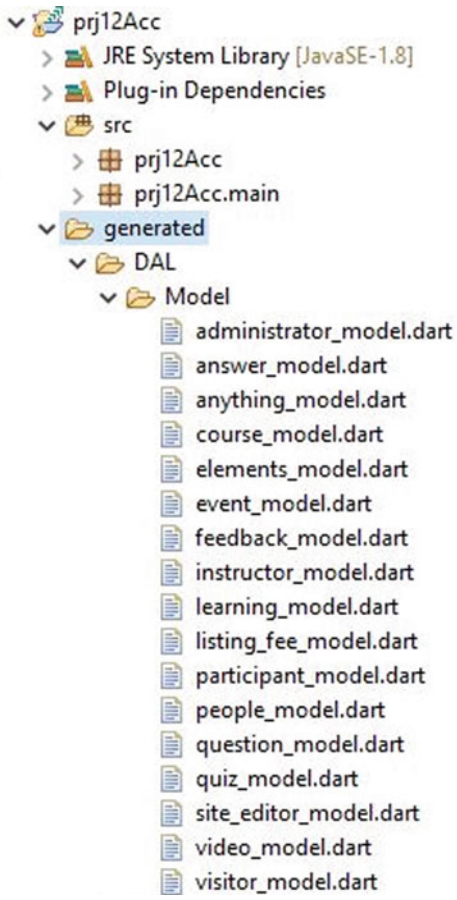


Fig. 9 Part of generated code of DAL layer

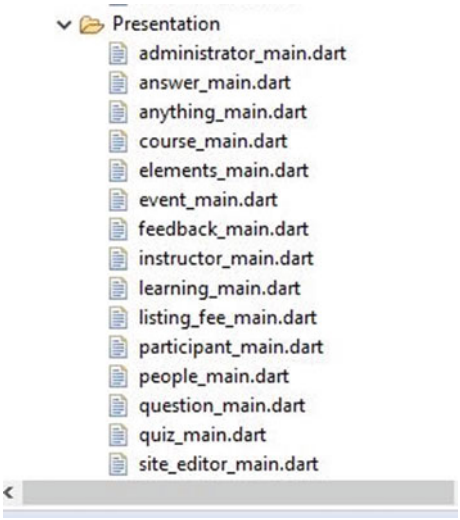


Fig. 10 Part of generated code in presentation layer

References

1. Palmieri M, Singh I, Cicchetti A (2012) Comparison of cross-platform mobile development tools. In: 16th international conference on intelligence in next generation networks, ICIN 2012, pp 179–186. https://doi.org/10.1109/ICIN.2012.6376023
2. Cohn M (2004) User stories applied: for agile software development. Addison-Wesley
3. Nasiri S, Rhazali Y, Lahmer M, Chenfour N (2020) Towards a generation of class diagram from user stories in agile methods. Procedia Comput Sci 170:831–837. https://doi.org/10.1016/j.procs.2020.03.148
4. Kraus A, Knapp A, Koch N (2007) Model-driven generation of web applications in UWE. In: CEUR workshop proceedings, vol 261, no Ist 016004
5. Rahmouni M, Mbarki S (2015) MDA-based modeling and transformation to generate N-tiers web models. J Softw 10(3):222–238. https://doi.org/10.17706/jsw.10.3.222-238
6. Castrejon JC, Lopez-Landa R, Lozano R (2011) Model2Roo: a model driven approach for web application development based on the Eclipse Modeling Framework and Spring Roo. In: CONIELECOMP 2011—21st international conference on electrical communications and computers, proceedings, pp 82–87. https://doi.org/10.1109/CONIELECOMP.2011.5749344
7. Daissaoui A (2010) Applying the MDA approach for the automatic generation of an MVC2 web application. In: 4th international conference on research challenges in information science—proceedings, RCIS 2010, pp 681–688. https://doi.org/10.1109/rcis.2010.5507266
8. Gharavi V, Mesbah A, Van Deursen A (2009) Modelling and generating AJAX applications: a model-driven approach. In: CEUR workshop proceedings, vol 445, pp 32–37
9. Huynh QT, Tran DD, Bui TMA, Le Nguyen P (2019) Development of rules and algorithms for model-driven code generator with UWE approach. Front Artif Intell Appl 318:531–544. https://doi.org/10.3233/FAIA190078
10. Bézivin J, Hammoudi S, Lopes D, Jouault F (2004) Applying MDA approach for web service platform. In: Proceedings—IEEE international enterprise distributed object computing workshop, EDOC, no 1, pp 58–70. https://doi.org/10.1109/EDOC.2004.1342505
11. Dimbisoa WG, Mahatody T (2018) Creating a metamodel of UI components in form of model independant of the platform, vol 6, pp 48–52
12. Deeba F, Kun S, Shaikh M, Dharejo FA, Hayat S, Suwansrikham P (2018) Data transformation of UML diagram by using model driven architecture. In: 3rd IEEE international conference on

cloud computing and big data analysis ICCCBDA 2018, pp 300–303. https://doi.org/10.1109/ICCCBDA.2018.8386531
13. Kherbouche M, Bouafia K, Molnar B (2019) Transformation of UML state machine to YAWL. In: Proceedings of 2019 IEEE 9th international conference on intelligent computing and information systems. ICICIS 2019, no Mdd, pp 215–220. https://doi.org/10.1109/ICICIS46948.2019.9014793
14. Miller J, Mukerji J (2003) MDA guide version 1.0.1. http://www.omg.org/docs/omg/03-06-01.pdf
15. OMG (2008) MDA. http://www.omg.org/mda
16. Appukuttan BK, Clark T, Reddy S, Tratt L, Venkatesh R (2003) A model driven approach to building implementable model transformations. Computer (Long Beach Calif) 1–14
17. Gerber A, Lawley M, Raymond K, Steel J, Wood A (2002) Transformation: the missing link of MDA. In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 2505, pp 90–105. https://doi.org/10.1007/3-540-45832-8_9
18. Tratt L (2006) The MT model transformation language. In: Proceedings of the 2006 ACM symposium on applied computing, Dijon, pp 1296–1303
19. https://www.eclipse.org/acceleo/
20. Sommer A, Krusche S (2013) Evaluation of cross-platform frameworks for mobile applications. In: Proceedings of the 1st European workshop on mobile engineering, p 365
21. Ribeiro A, Da Silva AR (2012) Survey on cross-platforms and languages for mobile apps. In: 2012 eighth international conference on the quality of information and communications technology (QUATIC), vol 255, no 260, 3–6 Sept 2012

Sentiment Analysis: Twitter Tweets Classification Using Machine Learning Approaches

Akula V. S. Siva Rama Rao, Sravana Sandhya Kuchu, Daivakrupa Thota, Venkata Sairam Chennam, and Haritha Yantrapragada

Abstract Analyzing public data from social media could give fascinating results and insights into the realm of public opinion on almost any product, service, or person. Sentiment analysis is a method for analyzing and interpreting Twitter data in order to determine public opinion. Creating sentiment analysis software is a method for measuring Twitter tweet perceptions programmatically. Results classify user's perspective via tweets into positive and negative, which is represented in result, may include polarity of the tweets. In this paper, we evaluate various machine learning algorithms performance to determine suitable algorithm for twitter datasets classification.

Keywords Sentiment analysis · SVM · Random forest and logical regression · Naïve Bayes Twitter tweets · Social media · Accuracy · Machine learning

1 Introduction

Twitter is a well-known Web-based media network for systems administration and miniature contributing to a blog, where the clients send and get messages as tweets. There are 199 million monetizable daily active users on Twitter (Twitter, 2021), with India accounting for 17.5 million of them. Today, social media has become main media as the primary means of creating and spreading public opinion [1]. Then,

A. V. S. Siva Rama Rao (✉) · S. S. Kuchu · D. Thota · V. S. Chennam · H. Yantrapragada
Department of Computer Science and Engineering, Sasi Institute of Technology & Engineering, Tadepalligudem, Andhra Pradesh, India
e-mail: shiva.akula@sasi.ac.in

S. S. Kuchu
e-mail: sravana5c1@sasi.ac.in

D. Thota
e-mail: daivakrupa5e9@sasi.ac.in

V. S. Chennam
e-mail: sriram5b4@sasi.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_39

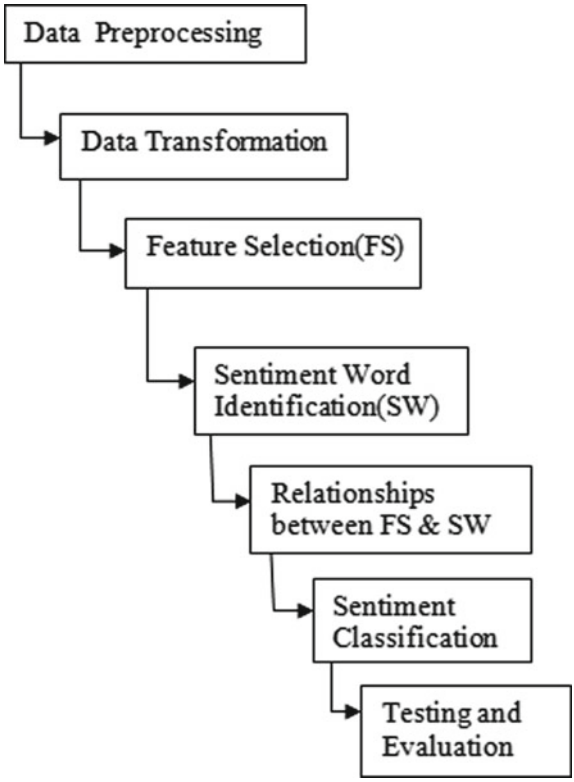


Fig. 1 Steps of sentiment analysis

again, in the securities exchange, new stories assume fundamental part on impacting the financial backers judgment and trust in stock worth [2]. Twitter goes about as media between individuals in the general public in trading the perspectives, often Twitter contributes extremely high in giving information to public [3] (Fig. 1).

Sentiment analysis is the way toward investing items surveys on the Web to know the general assessment or feeling with respect to that item [4]. Sentiment analysis is well known as a categorization task since it partitions a text orientation as positive, negative, or neutral. Perhaps, the most usually used ways to deal with notion classification are machine learning [5].

This paper includes classification of twitter tweets to determine the accuracy of the both positive and negative tweets utilizing various methodologies like support vector machine (SVM), random forest, logical regression, and Naïve Bayes.

2 Previous Work

Modern sentiment analysis researches are now being studied in a variety of subjects and languages. There are different ways that were used to classify Twitter data as part of the environment. Using emoji-based ideograms, the sentiment lexicon generates several types of emoticons. From Twitter data, a small number of standard emoticons were recognized, and the relevant emotions were identified. The sentiment analysis of this user-generated data is very useful in determining the public's opinion [6].

Wen, Song, and Tian presented another piece, the semantic and primary portion (S&S part), to identify stock costs utilizing support vector machine (SVM).

Wen, Xiao, and Gong explored different avenues regarding support vector machine and singular spectrum analysis (SSA).

Pak and Alexander proposed that the author creates a sentiment classifier using the corpus that can determine positive, neutral, and negative sentiments.

Agarwal and Apoorv described a popular tweeting site called Twitter and developed models for classifying "tweets" as positive, negative, or neutral.

Horakova and Marketa propose a technique that collects tweets from social media sites to provide a business intelligence view [7].

3 Methodology

To begin, we will sort the tweets into categories based on their content using support vector machine (SVM), logical regression, random forest, and Naïve Bayes. The tweets will be taken as strings and changed to numeric kind by utilizing TF-IDF.

3.1 Support Vector Machine (SVM)

The support vector machine calculation point l is to anticipate a hyper-plane in a N-dimensional space (N—the quantity of features) that recognizes information focuses. SVM is based on computational learning theory methods that try to reduce structural problems. The vectors that characterize this choice limit are named as help vectors [8].

3.2 Logistic Regression

For characterization issues, logistic regression is a directed machine learning method. The Supervised learning is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately. [9].

To break down the given text into an easily understandable format for the model, we will need to do some data preprocessing and cleaning [10].

The stages for preprocessing the tweets are as follows:

1. Lower Case—tweets are changed over to lowercase.
2. Hashtags—hashtags usually contain useful information, so they are a good choice. We are using the word hash instead of hashtag.
3. Detaching Twitter handles (@userid).
4. Removing special characters, numbers, and punctuations.
5. Tokenizing the tweets—the method of tokenizing a corpus is to break it down into its most fundamentally meaningful parts.
6. Stemming or lemmatization.

Model Building

To build the models, we will utilize logistic regression. It fits data to a logarithmic function to forecast the likelihood of an event occurring [11].

The following equation is used in logistic regression:

$$\ln(p/1 - p) = b0 + b1 * x \tag{1}$$

3.3 Random Forest Classifier

Random forests are effective methods. For regression and classification, random forests are applicable. By this process, significant feature can be identified. In Fig. 2, how a random forest would look like with two trees is illustrated [12].

3.4 Naive Bayes

For classification, the Naive Bayes approach is totally dependent on Bayes' theorem of probability [13]. That means, the algorithm basically expects that each input variable is irrelevant to the others. Making such a basic assumption about real-world data is very naïve [14].

If you use Naive Bayes for sentiment analysis, the algorithm will look at the individual words rather than the entire sentence if you say "I like Cricket." The meaning of words that are near to each other in a sentence is influenced by their position, and the position of words in a sentence is also important. However, sentences like "I like Cricket," "Cricket like I," and "Cricket I like" are all the same to the algorithm [15].

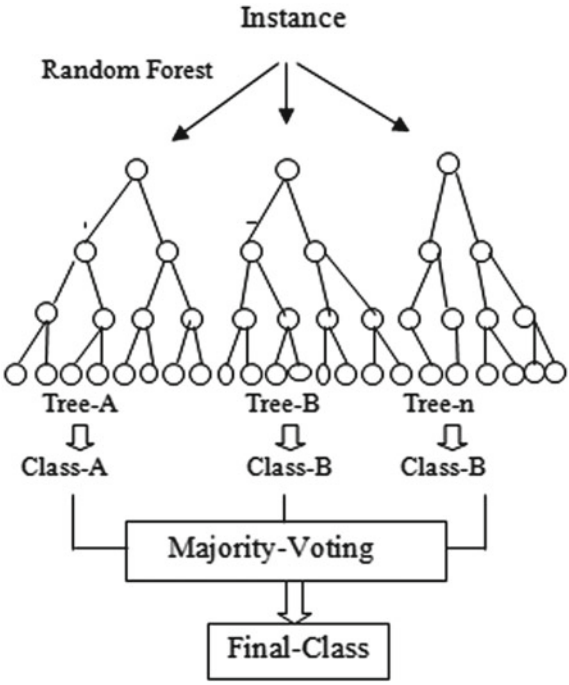


Fig. 2 Example for random forest

3.5 TF-IDF Vectorizer

The TF-IDF vectorizer technique executed is dependent on the frequency strategy; however, it is unlike the bag-of-words strategy in that it thinks about the occurrence of a word across the whole corpus, rather than simply in a solitary record.

3.6 Stemming Versus Lemmatization

Stemming: The way toward recognizing the morphological a root/base type of the word is known as stemming. The phrases "classical," "classify," and "classification" are conflated to the root word "classic." Similarly, "rectification," "rectify," and "rectified" are conflated to the stem word "recti" using a stemming algorithm [16].

Lemmatization: Lemmatization is like the stemming. The significant differentiation among stemming and lemmatization is that stemmer performed on a solitary word disregarding the specific circumstance, which created a stem that has various implications [17].

3.7 Feature Selection and Sentiment Word Identification

Feature selection is the process of reducing the number of input variables when developing a predictive model. To identify whether a word is a sentiment word, the most popular solution is to use a freely available sentiment lexicon such as HowNet or SentiWordNet [18].

Python programming was used to develop the algorithm. The used packages are mentioned below:

Pandas: Pandas is a Python package that can be used to analyze data that hasn't been labeled.

NumPy: The NumPy module is used to process multidimensional arrays, and the data can be imported into the code as data frames [19].

Re: Python programming supports for regular expressions that can be imported using the "re" module.

Nltk: The natural language toolkit, or NLTK for short, is a Python-based collection of modules and applications for symbolic and statistical natural language processing (NLP) in English.

3.8 Cleaning of Data

We begin by removing all of the tweets' special characters. Using the sub-function from package, regular experssion which remove all the single characters from left after removing the special character [20].

3.9 TF-IDF Vectorizer

The word frequency in a single text is defined as TF. Words and phrases can be used as terms.

4 Results

The work augmentation has been done to legitimize the exhibition of the current proposed strategy by contrasting it with the presentation of the every algorithm.

4.1 SVM

Support vector machine has an accuracy of 91.8% which is higher than both logistic regression and random forest algorithms.

4.2 Logistic Regression

The implementation of logistic regression was carried out with the help of a dataset of racist tweets obtained from Twitter, which had almost 15,000 observations. As a result, the method's accuracy was estimated to be around 79.5%.

4.3 Random Forest

Random forest implementation is carried out with the dataset consists around 14,000 tweets. The accuracy of this random forest classifier is 76%.

4.4 Naïve Bayes

The exhibition of the multinomial Naïve Bayes for different datasets is being imagined, which shows that the exactness of 69.6%.

Accuracy Depiction

In Table 1 and Fig. 3, the graphical representation of the exactness of the classifiers is being portrayed.

Table 1 Performance of algorithms

Sequence No.	Algorithm used	Accuracy (%)
1.	SVM	91.8
2.	Logistic regression	79.5
3.	Random forest	76.0
4.	Naive Bayes	69.6

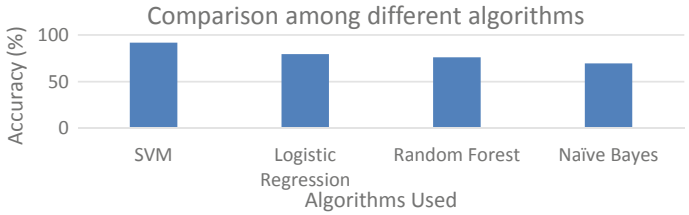


Fig. 3 Performance comparison among different algorithms

5 Conclusion

People's dependency on Internet evaluations is growing at a rapid rate, requiring the collection of diverse people's perspectives in order to provide a platform for learning what others believe. Sentiment analysis is not limited to a single dataset; it may be used to analyze any kind of review data. This paper contributes to this field of research by examining the performance of many sentiment classification algorithms and presenting a way for improving sentiment classification performance.

In traditional approaches, the SVM classifier gives the better accuracy when compared with remaining two algorithms.

6 Future Directions

In this paper, machine learning algorithms are considered; in future, deep learning and ensemble algorithms may be applied on different datasets.

References

1. Surya PPM, Seetha LV, Subbalakshmi B (2019) Analysis of user emotions and opinions using multinomial Naive Bayes classifier. IEEE, Coimbatore
2. Rathi M, Malik A, Varshney D, Sharma R, Mendiratta S (2018) Sentiment analysis of tweets using machine learning approach. IEEE
3. Bahrawi Bahrawi (Ministry of ICT Indonesia) (2019) Sentiment analysis using random forest algorithm online social media based
4. Sanchez-Rada JF, Iglesias CA (2019) Social context in sentiment analysis: formal definitions, overview of current trends and framework for comparison
5. Celiktug MF (2018) Twitter sentiment analysis, 3-way classification: positive, negative or neutral? IEEE
6. Ullah MA, Marium SM, Begum SA, Dipa NS (2020) An algorithm and method for sentiment analysis using the text and emoticon
7. Wang D, Zhao Y (2019) Using news to predict investor sentiment: based on SVM model
8. Hasan A, Moin S, Karim A, Shamshirband S (2018) Machine learning-based sentiment analysis for twitter accounts

9. Alsaeedi A, Khan MZ (2019) A study on sentiment analysis techniques of twitter data
10. Sailunaz K, Alhajj R (2019) Emotion and sentiment analysis from twitter text
11. Zhao Z, Zhu H, Xue Z, Liu Z, Tian J, Chua MCH, Liu M (2019) An image-text consistency driven multimodal sentiment analysis approachfor social media
12. Sentamilselvan K, Aneri D, Athithiy AC, Kani Kumar P (2020) Twitter sentiment analysis using machine learning techniques. Int J Eng Adv Technol 9
13. Kaur C, Sharma A (2020) Sentiment analysis of tweets on social issues using machine learning approach
14. Muhammad PF, Kusumaningrum R, Wibowo A (2020) Sentiment analysis using Word2vec and long short-term memory (LSTM) for Indonesia hotel reviews
15. Chaturvedi I, Cambria E, Welsc RE, Herrera F (2017) Distinguishing between facts and opinions for sentiment analysis: survey and challenges
16. Siva Rama Rao AVS, Ranjana P (2018) Machine learning based solution for homograph and auto-antonym ambiguities in social media corpora of sentiment analysis. J Adv Res Dyn Control Syst 10(3). ISSN 1943-023x. SCOPUS INDEXED JOURNAL
17. Siva Rama Rao AVS, Ranjana P (2020) Empower good governance with public assessed schemes by improved sentiment analysis accuracy. Electron Gov Int J 16(1/2). Inderscience Publishers. ISSN online: 1740-7508, ISSN print: 1740-7494. https://doi.org/10.1504/EG.2020.105252
18. Ansaia MZ, Aziza MB, Siddiqui MO, Mehraa H, Singh KP. Analysis of political sentiment
19. Ankita, Saleena N (2018) An ensemble classification system for twitter sentiment analysis
20. Drus Z, Khalid H (2019) Sentiment analysis in social media and ıts application: systematic literature review. Elsevier B.V.

Practical Fair Queuing Algorithm for Message Queue System

Le Hoang Nam, Phan Duy Hung, Bui Trong Vinh, and Vu Thu Diep

Abstract In microservices systems, the message queue is an architecture that provides asynchronous communication between services. The system has to receive requests from many sources to enqueue, while computing resources are limited. So, the queuing problem involves finding the order in which services are served fairly. Many theoretical studies have been conducted but no formal studies in the field of software development. This paper provides methods and designs of practical fair queuing algorithm in message queue of microservices system. Finally, the work presents the application of the design routines to a case study site.

Keywords Fair queuing algorithm · Scheduling algorithm · Message queue · Microservices

1 Introduction

Nowadays, applying microservices architecture has become more and more popular due to many advantages [1, 2]. Applications are decoupled into smaller, independent building blocks that are easier to develop, deploy, and maintain. Therefore, a request for processing is split into many steps; each will be handled by a service in the

L. H. Nam · P. D. Hung
FPT University, Hanoi, Vietnam
e-mail: nam19mse13037@fsb.edu.vn

P. D. Hung
e-mail: hungpd2@fe.edu.vn

B. T. Vinh
Hanoi Procuratorate University, Hanoi, Vietnam
e-mail: vinhbt@tks.edu.vn

V. T. Diep (✉)
Hanoi University of Science and Technology, Hanoi, Vietnam
e-mail: diep.vuthu@hust.edu.vn

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_40

system. This kind of design is useful, especially in statistics, analytics, and artificial intelligence (AI), where a single data processing request might be costly.

In the microservice architecture, the message queue is an almost indispensable component [3]. It provides communication and coordination for decoupled applications. All services have a common place for transporting information, instead of having to worry about receiving (sending) information from (to) others to continue processing data.

Looking at some systems, information can come from different sources such as users, reports, and data crawling with many types of data such as images, audio, and text. While computing resources are limited, some requests need to respond to requests as quickly as possible because processing time affects a company's revenue [4, 5]. Although microservices are designed to be easily scalable, it cannot be extended to infinity to handle all requests. Therefore, the queuing problems here are to find fairness order in which to serve these requests.

There have been some studies regarding the matter. In the paper "Queuing and Scheduling" [6], authors present the concept of "fair queuing." Fair queuing, or weighted fair queuing (commonly referred as bit-by-bit Round-Robin), is handlings N flows of arriving packets, which is problem of all other fair queuing problems. The idea of fair queuing is to scan all flows and to pick up a packet from each flow. But, with this scheme, larger packet will get more bandwidth, and it seems unfair about throughput. So, weighted fair queuing using estimating finish time of packet in the flow to determine number packets that will be picked up. Moreover, some flows can have higher priority than other, called as weights. A limitation of these original fair queuing is that scheduler need to scan all flows. If applying sorting algorithm here with more complex code, the complexity is $O(\log(N))$.

In the paper "Efficient Fair Queuing Using Deficit Round-Robin" [7], authors propose a new fair queuing algorithm, called Deficit Round-Robin (DRR). A scheduler handling N flows with quantum Qi, and a deficit counter (DC) for each flow. DRR scans for all non-empty flow at each round. When flow i is selected, its DC is incremented by its Qi, and after queuing a packet, the remaining, if any, is stored in its DC for the next round. The key idea is that at each round, the flow i can send at most value in its DC. So, a flow line can enqueue many packets with small size and fewer packets with large size. A flow with higher priority can be set up with higher quantum for enqueue more packets at a round. Therefore, the study proves that this scheme achieves nearly perfect fairness in terms of throughput, requires only $O(1)$ work to process a packet, compare with $O(\log(N))$ of other fair queuing algorithm. Deficit Round-Robin is also applicable to other scheduling problems where servicing cannot be broken up into smaller units.

Authors of the paper "Rotating Preference Queues: An Efficient Queue Management Scheme for Fair Bandwidth Allocation" [8] introduce rotating preference queues (RPQ). RPQ consists of a set of first-in-first-out (FIFO) output queue. These queues are dynamically rotated, and just one queue is used for dequeue and sending packets. The rotating of queues is to achieve fair bandwidth sharing, without cooperate with complicated buffer management scheme like Deficit Round-Robin. Thus, RPQ achieves fairness, requires $O(1)$ complexity to process a packet. One limitation

is that if both packets from same flow in a queue, the arriving packet will be discard. So, RPQ is only suitable for certain types of problems.

Other fair queuing algorithms [9–11] try to optimize based on original fair queuing and combine with other techniques to apply in other domains.

Although there are many studies as above, however, the works give solutions for some domains like packet scheduling, bandwidth allocation, and cloud computing but no implementation plan for the field of software development. This paper proposes a practical fair queuing algorithm for message queue system. This approach focuses on the ease and flexibility of implementation in software engineering.

The remainder of the paper is organized as follows: Section 2 describes the research methodology. Then, Sect. 3 provides the case study. Finally, conclusions and perspectives are made in Sect. 4.

2 Research Methodology

2.1 System Architecture

The system architecture is proposed in Fig. 1. Each source has many data with different size, need to be processed. Every interval time of Δt, the scheduler has a mission to pick up the number of input data to enqueue in the message queue. The

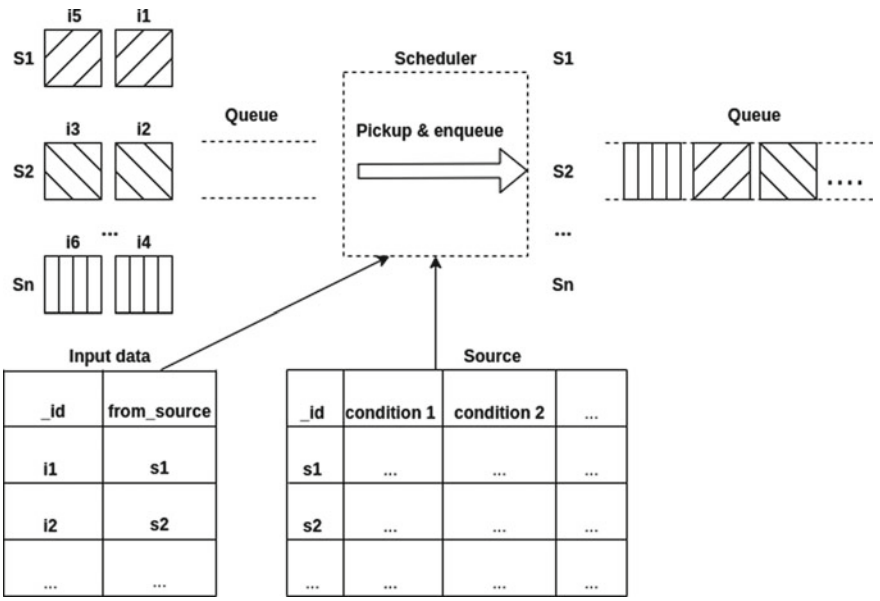


Fig. 1 Proposed design of system

problem here is to find the way that scheduler serving fairly with these sources, to make sure that no one waits too long to be served.

2.2 Proposed Fair Queuing Algorithm

This study uses ideas from fair queuing algorithms in [7] and makes it more practically by the context of software application. The number of sources here may be up to thousands or tens of thousands, even hundreds of thousands. And they don't always push data in their flows. So, if creating flows to store their arriving packets in the buffer, system needs to manage memory temporarily, make the scheduler stateful, hard to recover if system falls. The solution here is to store input data in database for long-term storage and making our service stateless. Inheriting the idea of deficit counter, each enqueue turn, the scheduler takes turns handling each source by get its condition (stored in table "Source") such as: max packet can be picked up, max total size of packet, and enqueue packets from table "Input data."

There are some advantages of this approach. Firstly, it can reduce complexity by indexing in database. Secondly, the fairness of enqueue, or the bandwidth of each source when enqueuing, is configured dynamically by many condition in database. Finally, by query of database, the number of sources on each enqueue turn reduces; thereby, the waiting time reduces.

3 Research Experiment

3.1 Case Study

The case study is a microservice system with the task of processing user input. There is a **Queue** which acts as message queue of the system. A single input of a user can be audio, image, or text, but here, they are collectively called *item.* When an item comes to the system, it is ready to be picked up, enqueued, and processed. The system has a large number of users, and each user also has many items waiting need to be served. Thus, there is a **Scheduler** with a mission to fair queuing those waiting items.

Every Δt time, **Scheduler** starts a turn to pick up items. The system has a database which consists of two tables: **Item** for storing information of items, and **User** stores information for fair queuing. These information include some conditions that help scheduler to determine number of items enqueued in each turn:

- *last_pickup* is the last time the scheduler selected items of an user. In each turn, a user with a **smaller** *last_pickup* will be prioritized first.
- *total_waiting* is the number of waiting items on each user input. In each turn, a user with a **larger** *total_waiting* will be prioritized first.

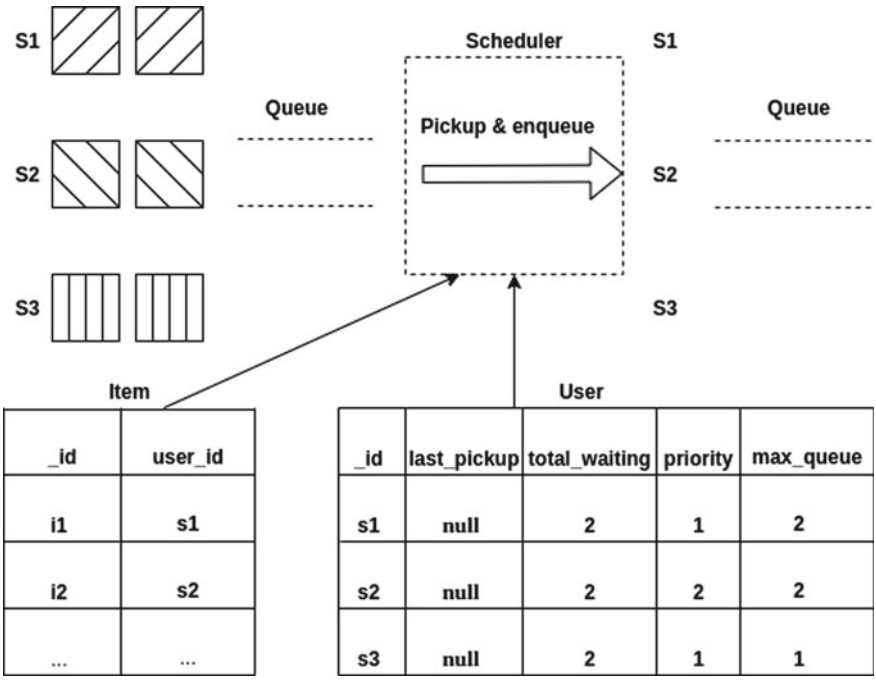
- *priority* is the number that has been determined for each user. In each turn, a user with a **larger** *priority* will be prioritized first.
- *max_queue* is a number that has been also determined for each user. In each turn, the number of selected items for a user must not be greater than this number.

Because here sorting on three fields (*last_pickup*, *total_waiting* and *priority*), the sort order is evaluated from left to right. So, users are first sorted by *last_pickup*; the users with the same *last_pickup* values are further sorted by *total_waiting*, and so on. The list of conditions and the sort order may vary to accommodate other cases.

3.2 Implementation of the Algorithm on a Case Study

The illustration of system on initial state is shown in Fig. 2. The example has three users, to ease the algorithm description.

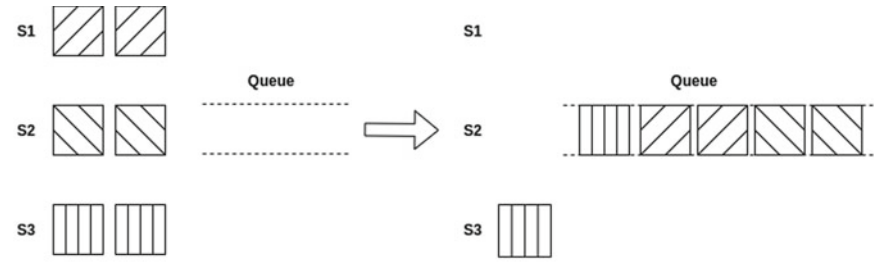
Figure 3 illustrates how the fair queuing algorithm works. After the first interval, at $t1 = \Delta t$, each user has two items on their flow. After querying with sorting by conditions in table **User**, a user list is selected in order as follows:



_id	user_id
i1	s1
i2	s2
...	...

_id	last_pickup	total_waiting	priority	max_queue
s1	null	2	1	2
s2	null	2	2	2
s3	null	2	1	1

Fig. 2 Initial state of case study



User

_id	last_pickup	total waiting	priority	max_queue
s1	Null	2	1	2
s2	Null	2	2	2
s3	Null	2	1	1

User

_id	last_pickup	total waiting	priority	max_queue
s1	Δt	0	1	2
s2	Δt	0	2	2
s3	Δt	1	1	1

Fig. 3 Pickup and enqueue at $t1 = \Delta t$

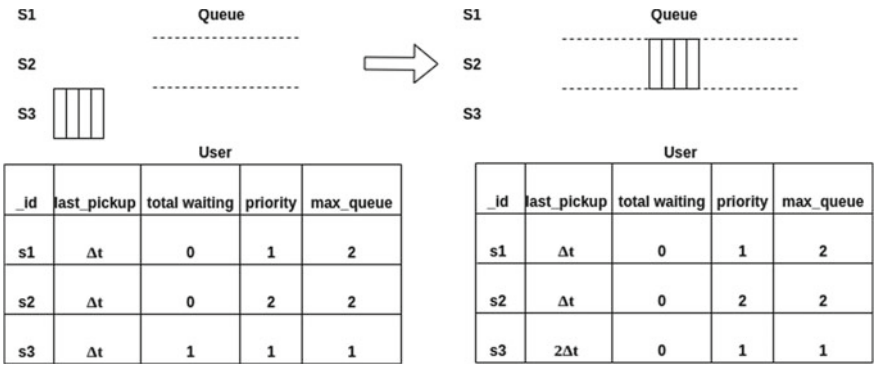
- $s2$: All users have same *last_pickup*, *total_waiting*, so $s2$ is chosen first because of its *priority*, with two items are enqueued.
- $s1$: $s1$ and $s3$ have same priorities so pickup in order according to the list. $s1$ is then chosen with two items are enqueued.
- $s3$: $s3$ is chosen last with 1 item is enqueued.

The above step is inspired from [7] to reduce the complexity of picking up an item. It is done by visiting each user and picking up their items, simultaneously in each turn, instead of scanning all users and choosing one like original fair queuing algorithm [6]. The max items queued of 1 user in each turn (*max_queue*) is also inspired by DC of [7], but the number can easily be changed when needed. Table **User** is updated after this turn, with updating *last_pickup* and *total_waiting*.

Figure 4 illustrates the system state at $t2 = 2 * \Delta t$. There is only $s3$ with *total_waiting* > 0, so table **User** sorts and choose 1 item of $s3$ to be enqueued. After that table **User** is updated with record of $s3$.

The removal of $s1$ and $s2$ which have no waiting items in this turn is an improvement of this algorithm compared to [7]. In [7], the DRR pointer reaches all users, even the empty ones. In contrast, by using the database query to obtain a list of users, the scheduler reduces offline users who have no waiting items like $s1$ and $s2$.

Figure 5 illustrates the system state at $t3 = 3 * \Delta t$. There are two users: $s1$ and $s2$ take precedence over $s3$ because of their *last_pickup* time is older than $s3$ ($\Delta t < 2 * \Delta t$). With $s1$ and $s2$, because $s1$ has more items waiting than $s2$, so, $s1$ items will be picked first. The final order is as follows: $s1$ with two items $\rightarrow$ $s2$ with one item $\rightarrow$ $s3$ with one item. After enqueuing, table **User** is updated in *last_pickup* and *total_waiting*.



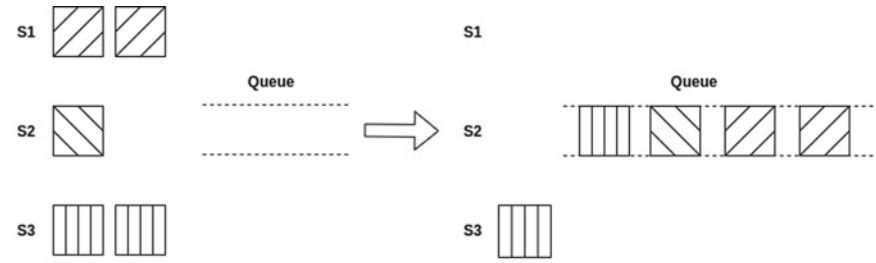
User

_id	last_pickup	total waiting	priority	max_queue
s1	Δt	0	1	2
s2	Δt	0	2	2
s3	Δt	1	1	1

User

_id	last_pickup	total waiting	priority	max_queue
s1	Δt	0	1	2
s2	Δt	0	2	2
s3	2Δt	0	1	1

Fig. 4 Pickup and enqueue at $t2 = 2 * \Delta t$



User

_id	last_pickup	total waiting	priority	max_queue
s1	Δt	2	1	2
s2	Δt	1	2	2
s3	2Δt	2	1	1

User

_id	last_pickup	total waiting	priority	max_queue
s1	3Δt	0	1	2
s2	3Δt	0	2	2
s3	3Δt	1	1	1

Fig. 5 Pickup and enqueue at $t3 = 3 * \Delta t$

3.3 Comparison to Other Works

The improvement of proposed algorithm is to reduce offline users with empty flows. With a software product, the number of offline users at a time is often quite large. So, reducing those users significantly reduces the waiting time for subsequent turns, reduces the I/O load of database, and makes the system works better.

In addition, in most implementations of previous fair queuing algorithms, the conditions for picking more items on some flows or precedence are predefined and difficult to change. But, with a software product, users may pay to have a higher priority or be served more. So, with the conditions store in database and can be changed, this system can easily meet the user's needs on flexibility and scalability.

4 Conclusions and Perspectives

The paper presents the application of fair queuing algorithm in a message queue system. The proposed algorithm focuses on the ease of implementation and flexibility in terms of "fair." Storing buffer data for enqueuing in database makes the scheduler service stateless, increases load capacity in case of many items of many users, and backups when scheduler service falls. The algorithm is based on querying and sorting techniques that are available on most current databases. Furthermore, the performance of querying is significant in terms of indexing in the database. Term "fair" of serving many users can be easily configured by update numbers in database. This "fair" is also very flexible and scalable because other can define their **Source** table with their predefined conditions, depends on their problems, and remove or add conditions easily. So, many problems of dealing with many inputs from many users may find it as a solution.

A limitation of this approach that is if the configuration is too biased on one or group of users, such as *max_queue* is large, or *priority* is considered first, the "fair" of enqueuing might not be guaranteed well.

In the future, applying this approach to a more complicated system is considered [12–14], and the optimizing of this approach will also be taken into account.

References

1. Fazio M, Celesti A, Ranjan R et al (2016) Open issues in scheduling microservices in the cloud. IEEE Cloud Comput 3:81–88. https://doi.org/10.1109/MCC.2016.112
2. Francesco PD, Lago P, Malavolta I (2018) Migrating towards microservice architectures: an industrial survey. In: Proceedings of the IEEE international conference on software architecture (ICSA), pp 29–38. https://doi.org/10.1109/ICSA.2018.00012
3. Message queues by IBM cloud education (2020) https://www.ibm.com/cloud/learn/message-queues
4. Top 5 ways system downtime can impact your company (2020) https://www.vxchnge.com/blog/ways-system-downtime-could-impact-company
5. Kumar P, Kalwani MU, Dada M (1997) The impact on waiting time guarantees on customers' waiting experiences. Mark Sci 16:295–314
6. Barreiros M, Lundqvist P (2015) QOS-enabled networks: tools and foundations, Chap 7. In: Queuing and scheduling, pp 117–142. https://doi.org/10.1002/9781119109136.ch7
7. Shreedhar M, Varghese G (1996) Efficient fair queuing using deficit round-robin. Proc IEEE/ACM Trans Netw 4:375–385. https://doi.org/10.1109/90.502236
8. Yang JP (2013) Rotating preference queues: an efficient queue management scheme for fair bandwidth allocation. IEEE Commun Lett 17:420–423. https://doi.org/10.1109/LCOMM.2012.121912.121840
9. Dekeris B, Adomkus T, Budnikas A (2006) Analysis of QoS assurance using weighted fair queuing (WFQ) scheduling discipline with low latency queue (LLQ). In: Proceedings of 28th international conference on information technology interfaces, Cavtat/Dubrovnik, pp 507–512
10. Patel Z, Dalal U (2016) Implementation and evaluation of dynamically weighted low complexity fair queuing (DWLC-FQ) algorithm for packet scheduling in WiMAX networks. China Commun 13:128–140. https://doi.org/10.1109/CC.2016.7489981

11. Dey S (2018) A secure fair queuing and SLA based slotted round robin load balancing approach for cloud data centers. In: Proceedings of the 3rd international conference on communication and electronics systems (ICCES), Coimbatore, pp 382–387
12. Hai MM, Hung PD (2020) Centralized access point for information system integration problems in large enterprises. In: Luo Y (ed) Cooperative design, visualization, and engineering. CDVE 2020. Lecture notes in computer science, vol 12341. Springer, Cham. https://doi.org/10.1007/978-3-030-60816-3_27
13. Tae CM, Hung PD (2020) A collaborative web application based on incident management framework for financial system. In: Luo Y (ed) Cooperative design, visualization, and engineering. CDVE 2020. Lecture notes in computer science, vol 12341. Springer, Cham. https://doi.org/10.1007/978-3-030-60816-3_32
14. Nam LH, Hung PD (2021) Building a big data oriented architecture for enterprise integration. In: Luo Y (ed) Cooperative design, visualization, and engineering. CDVE 2021. Lecture notes in computer science, vol 12983. Springer, Cham. https://doi.org/10.1007/978-3-030-88207-5_17

Indigenous Non-invasive Detection of Malnutrition-Induced Anemia from Patient-Sourced Photos Using Smart Phone App

K. Sujatha, N. P. G. Bhavani, U. Jayalatsumi, T. Kavitha, R. Krishnakumar, K. Senthil Kumar, A. Ganesan, A. Kalaivani, Rajeswary Hari, F. Antony Xavier Bronson, and Sk. Shafiya

Abstract The proposed online-based malnutrition-induced anemia detection smart phone app is built, to remotely measure and monitor the anemia and malnutrition in humans by using a non-invasive method. This painless method enables user-friendly measurements of human blood stream parameters like hemoglobin (Hb), iron, folic acid, and vitamin B12 by embedding intelligent image processing algorithms which will process the photos of the fingernails captured by the camera in the smart phone. This smart phone app extracts the color and shape of the fingernails, will classify the anemic and vitamin B12 deficiencies as onset, medieval, and chronic stage with

K. Sujatha (✉)
Department of EEE, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India
e-mail: drksujatha23@gmail.com

N. P. G. Bhavani
Department of ECE, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India

U. Jayalatsumi · T. Kavitha
Departmentt of ECE/Civil Engineering, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India

R. Krishnakumar
School of Electrical Engineering, Vels Institute of Science, Technology and Advanced Studies, Chennai, India

K. Senthil Kumar · Sk. Shafiya
Department of ECE, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India

A. Ganesan
Department of EEE, RRASE College of Engineering, Chennai, Tamil Nadu, India

A. Kalaivani
Department of CSE, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India

R. Hari · F. Antony Xavier Bronson
Department of Biotechnology, Dr. MGR Educational and Research Institute, Chennai, Tamil Nadu, India

Sk. Shafiya
Nalla Narasimha Reddy Group of Institutions, Hyderabad, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_41

specific and accurate measurements instantly. On the other dimension, this novel technology will place an end to the challenge involved in the disposal of biomedical waste, thereby offering a contactless measurement system during this pandemic Covid-19 situation.

Keywords Anemia · Malnutrition · Nail images · Machine learning · Smartphone app · Backpropagation algorithm · Discrete Fourier transform

1 Introduction

The primary objective behind this scheme is the non-invasive detection of hemoglobin (Hb), iron, folic acid, and vitamin B12 level in human blood. As on date, the blood samples are collected from the patients using invasive needles to measure the Hb, iron, folic acid, and vitamin B12 level by clinical examination. This is a painful procedure, and also, the disposal of biomedical waste becomes a challenging task. Current practice also has a disadvantage where there is a time gap involved in the analysis and increases the chance of risk due to infections. Telehealth monitoring using smart phone app enables continuous, safe, remote, and online diagnosis from the nail color and shape using intelligent image processing algorithms. The proposed scheme is a non-invasive method which uses, image processing and intelligent algorithms for online detection of anemia and malnutrition. Thus, the smartphone app once deployed will reduce the encumbrances in sampling the blood combined with reduction in cost needed for the development of a separate device or instrument thereby preventing the usage of very difficult and time-consuming laboratory-scale technologies. Also, this app can cater the needs of large number of patients enabling to maintain social distancing during this pandemic COVID-19.

2 Literature Review

The current practice proposed by Department of Electronics, Goa University, Goa Taleigao, India is the design and development of an algorithm which is used to evaluate the entire hemoglobin content in human blood. This scheme comprises of an embedded system which uses Arduino board. Currently, in hospitals, the monitoring and measurement of hemoglobin in patients are very significant so as to trace the high and low levels of Hb during surgery in operation theater, dialysis, donation of blood, during the testing of stress and intensive care unit (ICU) [1–3]. In all these places, the measurement of Hb level in the blood stream is done by an invasive approach where a sample of the blood is extracted from the patient. This system consists of a probe which is attached to the finger with one LED attached for each finger on a board with multiple chips where the wavelengths vary between 670 and 950 nm, which

is built on a single silicon photo detector. Photoplethysmography (PPG) principle-based instruments with a trans-impedance amplifier are constructed to gather the signals for the finger. This University at Goa has identified a non-invasive to monitor the level of the Hb without pain and infection-less devices and instruments. This method not only facilitates a real-time monitoring but also produced outputs which are of comparable standards. This researcher has taken totally eight wavelengths which are compared with the standard values using invasive methods. The outputs are tested and then compared with the standard values obtained in the laboratory. This method is found to show increased accuracy in the measurement of hemoglobin, with minimum deviation between the results obtained from the embedded system model and laboratory results. The error difference is estimated to be 1.3778–0.3499 g/dL, and regression coefficient ($R^2 = 0.964$) is obtained during Hb analysis [4, 5].

This proposed scheme does not require any device apart from a mobile phone which almost, all the people have with them, as it forms a part of their daily need. So, this method as compared with the existing technology eliminates the usage of a separate device for measuring anemia, and it does not give any information about malnutrition [6, 7]. The accuracy of the results obtained using this smartphone app will be a smart, optimal, and cost-effective one as compared with the device or instrument which was developed using embedded system. Even though the device is a compact one and cost less, it exists as a separate instrument which the people will need to procure if the person is affected by anemia. On the other hand, this proposed smartphone app using intelligent image processing algorithms can be downloaded by anybody whoever possess a smart phone as it has usage of smart phones has become a part and parcel of life nowadays during this COVID-19. This method serves as an effective layman approach for measurement of Hb, iron, folic acid, and vitamin B12 avoiding large gathering at testing laboratories, thereby enabling social distancing during this pandemic situation [6, 7]. A comparative analysis is depicted in Table 1.

3 Objectives

The main objectives are as follows:

i. To implement a non-invasive system for detection nutrition deficient anemia by screening of hemoglobin iron, folic acid, and vitamin B12 levels from the nail color and shape.
ii. To relieve the patients from pain that is introduced during the collection of blood samples using needle and syringe for laboratory analysis.
iii. To minimize the cost involved in laboratory analysis which will require the usage of chemicals and reagents.
iv. To eliminate the biochemical hazards involved in disposal of related biomedical waste.

Table 1 Comparative analysis for anemia diagnostic methods with the app

Device	Non-invasive/invasive technique	Smartphone app technology	Stand-alone device (apart from mobile phone)	Cost-effectiveness
Complete blood count hemoglobin levels (gold standard)	Invasive	No	No	No
Hemocue	Invasive	No	No	No
Masimo co-oximetry	Non-invasive	No	No	No
Conjunctival analysis	Non-invasive	Yes	No	Yes
WHO color scale	Invasive	No	No	Yes
HemaApp	Non-invasive	Yes	No	Yes
Smartphone anemia app	Non-invasive	Yes	Yes	Yes

v. To nullify the time delay involved in laboratory-scale blood sample analysis to identify anemia, due to the lack of micronutrients such as iron, folic acid, and v itamin B12.
vi. To enable a remote, online, and continuous monitoring system to screen the nutritionally deficient anemia at the earlier stage which can be corrected by supplementing the iron, folic acid, and vitamin B12 to prevent further complications by analyzing at least 70,000 samples within no matter of time.
vii. To develop indigenous intelligent image processing algorithms and a smartphone-based app, this can accommodate many users thereby enabling social distancing during this pandemic COVID-19.

4 Novelty

The project's research highlight is the development of smart phone app which uses image processing algorithms like Fourier coefficients using discrete Fourier transform (DFT) for feature extraction, where these Fourier coefficients are found to show variation with respect to various stages of anemia (onset, medieval, and acute stages) along with the measurement of Hb, iron, folic acid, and vitamin B12 and invariably classified or clustered using intelligent algorithms like backpropagation algorithm (BPA) using two different types of activation function (sigmoid and threshold). These images will be processed by two algorithms: One is the classification algorithm, and

the other one is the clustering algorithm, which is attempted, to evaluate the detection efficacy during diagnosis.

The features (Fourier coefficients) are extracted from the images of the fingernails, whose photo will be captured using the camera in the smart phone. The motivation of the proposed method is to develop a non-invasive, contactless, online, continuous, and cost-effective smart phone app for detecting anemia and malnutrition. This kind of smartphone app which relies on intelligent image processing algorithms can examine thousands and thousands of patients (samples) in a very short duration of time with high precision and accuracy, thereby resolving time complexity, amidst the pandemic, Covid-19. This indigenous system will offer an effective detection system and is user-friendly that even a layman possessing a smart phone will be capable of self-oriented diagnosis without the help of a clinician.

5 Methodology

Preliminary technology is proposed here, which can identify anemia and malnutrition using a non-invasive, self-screening tool for use by the common people. The sample images of the fingernails are captured using smart phone. These images are then transferred to the laptop where it is preprocessed for noise removal using median filter. Followed by de-noising, the features like Fourier coefficients are extracted along with shape descriptors which are used as inputs to train, test, and validate the backpropagation algorithm (BPA). The images of the fingernails from smartphone were transferred to a laptop computer to remove the noise if it is present. Features corresponding to shape and color of the fingernails were extracted from the samples collected. An automatic selection algorithm is used to identify the region of interest (RoI) from which finger nail color and shape were extracted. The thumb finger is excluded, and RoI is selected from each finger with 786 pixels. Nail color data were taken out from RoI and averaged for each finger.

Anemia and malnutrition screening amid the people is to construct an algorithm and develop a smart phone App to screen the anemia, iron deficiency, folic acid, and vitamin B12 for three categories (mild, moderate, and severe) along with the normal range values for the entire study population with mild anemia, mild iron deficiency, mild folic acid levels, and mild vitamin B12 (18 samples), moderate anemia, moderate iron deficiency, moderate folic acid levels, and moderate vitamin B12 (20 samples), severe anemia, severe iron deficiency, severe folic acid levels, and severe vitamin B12 (13 samples), and finally, non-anemia, normal iron level, normal folic acid levels, and normal vitamin B12 levels (51 samples), respectively. The developed algorithm is also tested and validated by taking new set of samples pertaining to the above said category. These results are once again cross-validated with the laboratory-scale results. The captured nail images are preprocessed for noise removal. Histogram analysis is done to identify the variation with respect to the nail color in coordination with the Hb, iron, folic acid, and vitamin B12. Subsequently, features are extracted using DFT, and the Fourier coefficients are used as inputs to

Table 2 Laboratory test values for anemia, iron, folic acid, and vitamin B12

Types of anemia	Hemoglobin level	Folic acid level	Serum iron level	Vitamin B12
Normal persons	Male 13 g/dl Female 12 g/dl	Adults 2–17 ng/mL Children 5–21 ng/mL	65–175 mcg/dL for men 50–170 mcg/dL for women 50–120 mcg/dL for children	**Normal**. The normal range for vitamin B12 in the blood is between 200 and 900 nanograms per milliliter (ng/mL)
Mild anemia or onset	9–11 g/dL	Adults 2–17 ng/mL Children 5–21 ng/mL	Less than the normal value	Less than the normal value
Moderate or medieval anemia	7–9 g/dL	Less than the normal value	Less than the normal value	Less than the normal value
Severe or chronic anemia	Less than 7 mg/dl	Less than the normal value	Less than the normal value	Less than the normal value

train the ANN classifier which uses BPA for estimation of the Hb, iron, folic acid, and vitamin B12. This procedure is implemented using MATLAB simulation studies at our organization (Table 2).

6 Results and Discussion

Laboratory-scale testing for anemia and malnutrition is done.

The estimation of hemoglobin was performed using cynomethemoglobin method (Drabkin D, Austin).

1. The morphology of cells was analyzed using hematology analyzer.
2. Estimation of iron in the samples was performed by the ammonium thiocyanate method using kit procedure.
3. The estimation of folic acid and vitamin B12 is done by using ELISA kit procedure.

6.1 Nail Image Acquisition

The nail images are captured using the camera in Apple 7S smart phone. During this process, the intensity of the light is increased so as to enhance the quality of the images captured using Lux meter app which is also installed on the same smart phone. Nearly, 700 pixels from the captured nail image are cropped and considered for further analysis.

6.2 Image Preprocessing

The preprocessing of the nail images includes two parts. The first one is the histogram analysis which gives the frequency of occurrence of various pixel intensities. This analysis is done to identify whether there exists a variation in the value and pattern of the pixels due to color variation in the nails which in turn depict the deficiencies of Hb, iron, folic acid, and vitamin B12. The second stage of preprocessing is the removal of noise which occurs due to variation in illumination during the capturing of nail images. This can be eliminated by using DFT which will also act as a filter to remove any noise present.

6.3 Feature Extraction

Features are the basic building block of an image which gets repeated in various directions. The features are unique values of pixels which can be mathematically modified to infer useful information. The discrete value of the Fourier coefficient which gives the power spectrum value is used for further analysis.

The FFNN trained with BPA is used for categorizing the nail images from which the values of Hb, iron, folic acid, and vitamin B12 are estimated for various conditions like normal, mild, moderate, and severe anemic and malnutrition conditions are shown in Fig. 2. The BPA algorithm will be using two types of activation function (namely sigmoid and threshold), to identify which type of activation function will infer fast convergence with least value of the error. The simulation results using MATLAB for malnutrition-induced anemia is illustrated in Figs. 1, 2, 3, 4, and 5, respectively.

The simulation results illustrate that the designed algorithm is capable of estimation of malnutrition-induced anemia with an error of 0.0001, and classification efficiency is inferred to be 97% accurate.

The following block diagram explains a high-level architecture of the mobile application (to be developed). This architecture will have the following

Mobile Client—this will act as a user interface and can be installed on their mobile phones. Server—deployed on AWS as an API, will be responsible for the data processing and notifications.

The following are the various components present on the mobile application

Mobile application is the container which enables the user to authenticate and capture his/her nails. This application will consume the various APIs developed, to process the captured images and detail out the deficiencies based on the algorithm.

Server APIs are the connectors which bridge the mobile application and the data processing engine hosted on the application server. These API utilize its own authentication mechanisms for a secured data transport. Data processing engine is developed with a pre-approved algorithm to process the captured images, match it with the prebuilt models, and provide a response back with the details. This module also has its

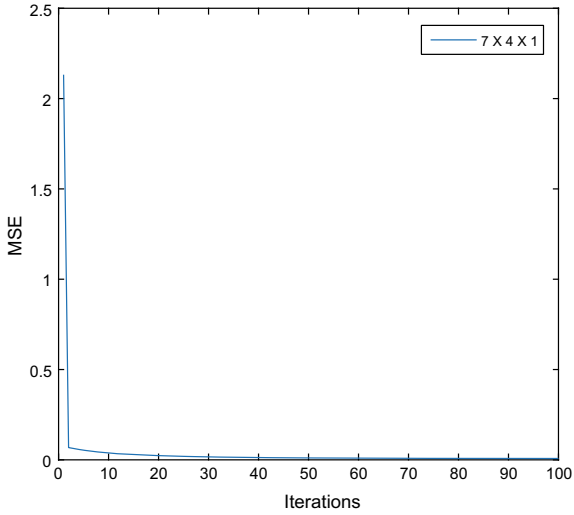


Fig. 1 Output for iterations versus MSE during training of BPA

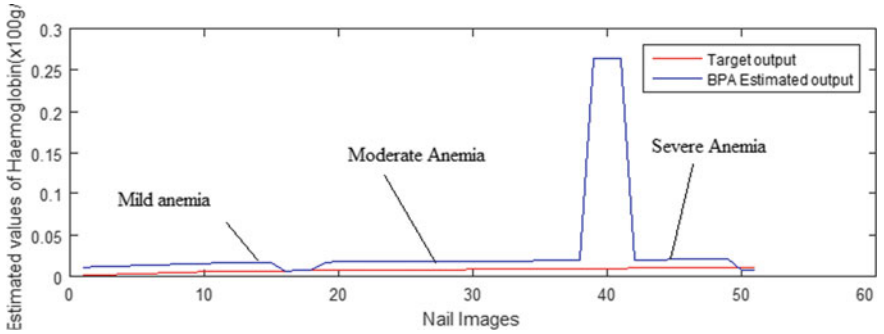


Fig. 2 Hemoglobin estimation during anemia detection

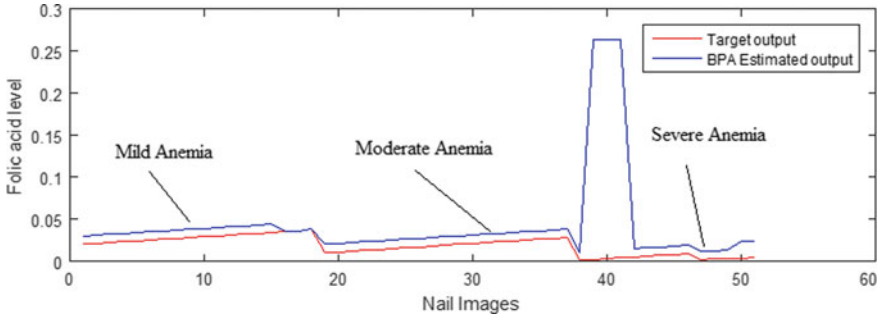


Fig. 3 Serum iron estimation during anemia diagnosis

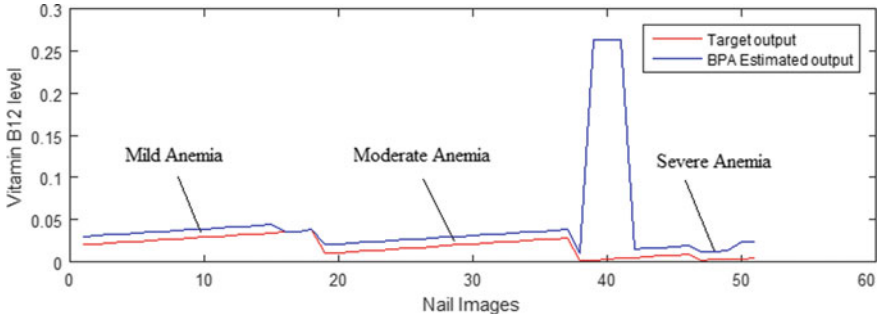


Fig. 4 Vitamin B12 estimation during anemia diagnosis

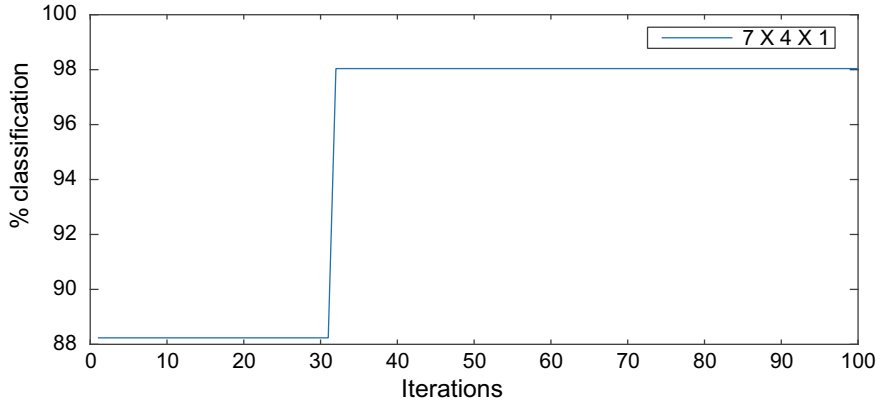


Fig. 5 Classification for diagnosis of malnutrition-induced anemia

data ingestion module for the administrator to upload more samples or tweak sample results. Error handling module provides the base framework to capture erroneous data or any server-related downtime. This module will be responsible for providing the user-readable messages.

7 Conclusion

As a part of the future scope of this work, the other parameters like vitamin A, ferritin, and the left-over constituents of the blood can also be analyzed and estimated by using this smart phone app, where the nail images are captured and analyzed. The RBC, white blood corpuscles (WBCs), and platelet count can also be obtained using this smart phone app which captures the nail images and uses a smart phone app with inbuilt intelligent image processing algorithms so that it is also possible to diagnose

diseases like dengue, leukemia, and leukopenia using a non-invasive method. Even the bone density can be estimated by detecting the calcium levels present in the blood for normal conditions as well as mild, moderate, and severe deficiencies; as calcium deficiency leads to osteoporosis. So, this kind of smart phone app can serve as an easy tool to detect osteoporosis without performing dual-energy X-ray absorptiometry. This smart phone-based app can also be extended to identify the presence of corona virus by measuring the physiological parameters of the blood from the nail color using an improvised version of the smart phone-based app.

References

1. Vindhya J, Nath A, Murthy GVS, Metgud C, Sheeba B, Shubhashree V, Srinivas P (2019) Prevalence and risk factors of anemia among pregnant women attending a public-sector hospital in Bangalore, South India. J Fam Med Prim Care 8(1):37–43
2. Government of India (2016) National family health survey 4-key indicators: 2015–2016. Natl Fam Health Surv 4:1–8
3. Orkin SH (2000) Diversification of haematopoietic stem cells to specific lineages. Nat Rev Genet 1:57–64
4. Rati SA, Jawadagi S (2012) Prevalence of anemia among adolescent girls studying in selected schools. Int J Sci Res 3(8):1237–1242
5. Premalatha S, Valarmathi T, Parameshwari S, Sundar S, Kalpana J (2012) Prevalence of anemia and its associated factors among adolescent school girls in Chennai. Epidemiol Open Access 12(2):2–5
6. Jadhav BN (2016) A critical review on iron deficiency anaemia in female population of developing India. Int J Fauna Biol Stud 3(5):116–119
7. Kraitl J, Ewald H, Gehring H (2005) An optical device to measure blood components by a photoplethysmographic method. J Opt A Pure Appl Opt 7(6):S318–S324

Speaker Diarization and BERT-Based Model for Question Set Generation from Video Lectures

Sravanthi Nittala, Pooja Agarwal, R. Vishnu, and Sahana Shanbhag

Abstract The year 2020 and the onset of the pandemic has, by and large, rendered the traditional classroom-based learning experiences obsolete. This rapid change in the learning experience has brought with it the opportunity to explore new avenues such as online learning. In this paper, we outline a methodology that aims to aid this paradigm shift by proposing an automated question set generation model based on the video lectures delivered by the teacher. We study the usage of pre-trained Scientific Bidirectional Encoder Representations from Transformers (SCIBERT) checkpoints for question generation on text derived from video lectures. The proposed methodology takes into consideration the presence of multiple speakers in the video lectures and employs speaker diarization for obtaining the audio corresponding to the teacher's voice. Diarization is implemented using the mean shift clustering algorithm. For the purpose of answer-agnostic question generation, pre-trained SCIBERT checkpoint is used to warm start an encoder–decoder model. The model is fine-tuned using the SQuAD dataset. The results show that the model is able to successfully generate questions for a given context derived from the transcript of a diarized video lecture.

Keywords Speaker diarization · BERT · Question generation · NLP · SCIBERT · Mean shift algorithm

1 Introduction

Automated Question Generation is the process of training a model, or providing rules using which, questions can be automatically generated from a given context by a system. This process plays an important in many environments, not limited to the field

S. Nittala (✉) · P. Agarwal · R. Vishnu · S. Shanbhag
PES University EC Campus, Bengaluru, India
e-mail: ndpsravanthi@gmail.com

P. Agarwal
e-mail: poojaagarwal@pes.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_42

of education [24]. The combination of automated question generation with speaker diarization allows the system to be used for generating questions from video lectures. Such systems prove useful in easing the burden of both students and educators, especially while navigating the relatively unfamiliar territory of online learning that has become a necessity in many countries due to the COVID-19 pandemic. The process of automated question generation is not a new concept, with earlier systems making use of syntactic and semantic rules for the purpose of identifying answers and generating different types of questions such as gap-fill questions, multiple choice questions, and "wh-" questions [10, 15]. With the advancements of technology, computer scientists are now making use of Deep Learning and Machine Learning models for the same purpose, transitioning to what is referred to as neural question generation [5, 20, 23, 25].

The recent increase in the usage of pre-trained models in the field of NLP for processes such as extractive summarization [13, 19], question answering [16] has led to more exploration of the potential such models hold in other aspects of NLP. Our proposed methodology uses the SCIBERT [1] model for question generation. The SCIBERT model is based on the Bidirectional Encoder Representation from Transformers (BERT) model. BERT [4] is a pre-trained model developed by Google for the purpose of Natural Language Programming. The model comprises an encoder stack with multiple attention heads and is commonly used to generate intermediate representations of input sentences. SCIBERT shares the architecture of BERT but is trained on a corpus of scientific papers in order to accommodate NLP tasks in the scientific domain. In the context of this project, the SCIBERT checkpoint is used to initialize an encoder–decoder network, which is in turn fine-tuned for the purpose of generating questions from the given context, particularly for those contexts belonging to the domain of computer science.

An important component of our project is the speaker diarization function. Speaker diarization is a process that is used to identify the speakers in a given audio clipping. Speaker diarization has been applied for many purposes such as identification of dialog between Air Traffic Control and pilots [7], conversational telephone speech [12] and to more recent applications in the medical field for identifying speech patterns in autistic patients [9]. Since, the project generates questions from the context of subject lectures delivered via video; it becomes essential to consider only the speech corresponding to the teacher to ensure factually correct data is input to the subsequent steps. This introduces the need to diarize the audio obtained from the lectures. Diarization as a process can be broadly described as having two steps, extracting speech segments and classifying them into the different speakers. The process of classification can be done using clustering algorithms [9], algorithms such as the Viterbi algorithm have also been employed as shown in [2].

The proposed system makes use of the unsupervised clustering approach for the process of speaker diarization. Once the diarized audio clippings are obtained, they are converted into text by making use of speech-to-text API which produces the required transcript. The obtained transcript is processed by the SCIBERT-based question generation model and the output is given in the form of questions based on the

given context. The answers to the produced questions are then used to produce distractors for Multiple Choice Questions using the word2vec [8] library. The question generation model is trained on the Stanford Question Answering Dataset (SQuAD) [18].

2 Related Work

The proposed methodology consists of the following steps: speaker diarization, speech-to-text conversion, question set generation, and distractor generation. There has been an immense amount of research that has been carried out in each of the individual fields, and in some cases, a combination of two or more of the fields.

2.1 Speaker Diarization

Speaker diarization has its roots based in the usage of unsupervised clustering algorithms [12], but recent research has seen a shift towards usage of neural networks and deep learning in the diarization process [6]. Fujita et al. [6], for example, used a single step through the neural network to obtain the diarized output, moving away from the usual method of having an extraction step followed by a classification step, by modelling the problem as a multi-label classification problem. It was implemented using a Bidirectional Long-Short Term Memory (BLSTM) network, and outperformed the existing clustering-based models. It performed well on overlapped speech segments as well.

Huang et al. [11] proposed a methodology for speaker diarization by making use of a Region Proposal Network (RPN) to propose the speech segment boundaries from within the audio signals given as input to the convoluted layers of the network. These speech segments were then processed for classification. They considered making use of the k-means clustering algorithm for classification and the speech segments are classified according to the speaker. This paper makes use of RPN to take care of overlapped speech as well.

Gorodetski et al. [9] mainly focused on recognizing the speech of a child and their therapist in recordings of therapy sessions for autistic children. They make use of the Viterbi algorithm for the classification of the speech segments. The authors have taken into consideration the frequency and pitch differences between the child, and therapist. The recordings were segmented and manually labelled depending on the speaker and then used to train a Gaussian Matrix Model using the Baum–Welch algorithm, which along with a recognition network, was used for the Viterbi algorithm for diarization.

2.2 Question Generation

For the process of Question Generation, there are a number of implementations including using syntactic, and semantic rules to generate questions or using NLP pipelines [5, 10, 15]. The papers reviewed below, use neural question generation processes. Du et al. [5], was one of the first works to move away from, and outperform the previously used rule-based models or NLP pipelines, by making use of sequence-to-sequence learning instead. The model used a recurrent neural network encoder–decoder architecture and global attention mechanism. The encoding is done at both paragraph level and sentence level, where the latter is concatenated to the former before being passed to the decoder. Encoding is done using a bidirectional LSTM at both sentence and paragraph level. For training, the SQuAD dataset [18] is used consisting of sentence–question pair with inference after training done using beam search.

Scialom et al. [20] highlights the difference between answer-aware and answer-agnostic question generation, while developing a model using transformers for the latter purpose. Answer-agnostic QG refers to that method of QG where the answer is not provided along with the context as input. The input consists of only the context from which the question is to be generated. The authors make use of transformers and train it on a crowd-worked SQuAD dataset [18] which consists of (sentence, question) pairs. In order to handle rare or unseen words, the authors have proposed three methods besides the context-free word representations offered by transformers. These methods are: placeholder strategy, copying mechanism, and contextual embedding.

The paper by Wu et al. [23], implements answer-agnostic question generation by making use of Sequence-to-sequence architecture. The implementation is such that, first the possible question type is predicted using the last hidden state of the encoder following which the question is generated. Additionally, the authors also propose a copy loss function for an enhanced copy mechanism. The probability distribution of the question types is then generated and the top question type is taken as input to the decoder along with the last encoded hidden state to generate the questions.

Chang et al. [2] recognizing the need for question generation from long sequences of input, make use of the BERT model [4]. There are three implementations considered: first, the BERT model is used as is for the purpose of question generation, second, it is implemented as a sequential model where the decoded output depends on the previous output (BERT-SQG), and finally, the sequential model is considered once again, this time the answers are highlighted within the context using the HL token (BERT-HLSQG).

Ma et al. [14] proposed a template-based method to extract questions from MOOC (Massive Open Online Courses) video lectures. The model generates the questions from the video subtitle; it does so by first extracting facts and representing them as a knowledge graph. The knowledge graph then serves as an input to the template-based model which considers each fact and the relations between the entities of the facts to generate the questions accordingly. The model was able to successfully generate questions in domain-specific MOOC videos.

Wang et al. [22] implemented question generation from educational textbooks by making use of recurrent neural networks. The model generates factual questions from source context, where the answers were highlighted in the context given as input. They made use of pointer network which allows certain parts of the context to be given more importance. The model, however, does not scale well and is limited to generating only factual questions. The authors found that even after training on a general-purpose questions dataset, the model performed well on domain-specific textbook source contexts.

The expanding number of applications of pre-trained models such as BERT [4], GPT-2 [17] etc. In the field of NLP, led to research into using these models to warm start encoder–decoder models by Rothe et al. [19]. The authors explored the results of using publicly available pre-trained checkpoints to initialize a Transformer-based Seq2Seq network in various text generation tasks. They made use of checkpoints of BERT [4], GPT-2 and RoBERTa for the purposes of abstractive summarization, machine translation, split and rephrase, and sentence fusion. The authors found that the models gave promising results when the encoders were initialized using the pre-trained checkpoints.

The model proposed by Srinivas et al. [21] for generating distractors does so for questions that consist of the code of a function, with answer being an appropriate description of the function. A database consisting of pairs of the form (C, D), where C is the function code and D is the corresponding function description, is used. The tool picks pairs of the form (C', D') which are similar to (C, D) using Abstract Syntax Tree (AST) and doc2vec, which are then ranked based on the similarity in description on being compared to D and the top 5 ranked values are taken as distractors.

3 Motivation of the Work

The proposed model aims to generate questions from context obtained as transcripts of video lectures. The main motivation behind this work includes:

- To provide a convenient method by which educators can aid the learning of their students without adding to their own workload.
- Proposing a new methodology to generate questions from video lectures for which the existing work is limited as outlined in the survey by Das et al. [3].
- Verifying the usability of the mean shift algorithm for the purpose of diarization, for which, to our knowledge, it has not been used in existing work.
- Testing the efficiency of initializing encoder–decoder networks with SCIBERT [1] checkpoints to generate answer-agnostic questions.
- Developing a question generation system that works in an answer-agnostic environment.

4 Terminologies Used

- Diarization—Diarization is the process of dividing an audio recording into segments and identifying the speakers within the resultant segments.
- Word2Vec—Word2Vec [8] makes use of neural network model to learn the relationship between words from a large corpus of text. Here, it is utilized for generating distractors.
- Distractors—Distractors are the incorrect options in multiple choice questions. These serve as one of the outputs of the proposed system.
- Sequence-to-sequence model—A Seq2Seq model is one that takes a sequence of tokens as input and generates a sequence as output. This model is used in this system for question generation.
- Bidirectional Encoder Representations from Transformers—BERT [4] is a pre-trained machine learning model for natural language processing (NLP) based on transformers developed by Google. The pre-trained SCIBERT model is used to warm start the Seq2Seq network.
- Warm-starting—Warm-starting an encoder–decoder involves initializing the parameters of an encoder or decoder network such as the weights by making use of the values in a pre-trained model such as BERT, GPT-2, RoBERTa, etc.

5 Proposed Methodology

Based on the literature survey done, and the existing systems, we propose a system for generating questions, given a video lecture as input. Speaker diarization is carried out to obtain the speech segments corresponding to the teacher's voice from the lecture's audio. The implementation for this follows the same pattern of segmentation followed by clustering, with the difference coming in terms of the clustering algorithm used. In order to account for the fact that the number of speakers will not be known beforehand, we propose to use the mean shift algorithm for the purpose of clustering. The clustering factor, in our context, is the frequency of the audio at a given time. The largest cluster obtained is taken as that belonging to the teacher's voice. From timestamps associated with each segment in the cluster, we split the original audio/video file into multiple smaller files containing single dialogue speech segments. The obtained speech segments are then converted to text using the Google speech-to-text library.

The question generation is implemented by using pre-trained SCIBERT [1] checkpoints in an encoder–decoder network. The inputs to the encoder and decoder are the context paragraph and the question to be generated, respectively. The training dataset is formatted such that the context or the paragraph from which the questions are to derived, is given as input to the encoder after tokenization by the BERT tokenizer initialized with SCIBERT [1]. The tokens are generated along with the [CLS] and [SEP] tokens at the beginning and end, respectively. The input to the encoder is in the following form, where $p1$, $p2 \ldots pn$ are the tokens generated for the paragraph,

$$[\text{CLS}], p_1, p_2, \ldots, p_n, [\text{SEP}]$$

The encoder generates the word embeddings according to the weights initialized by the SCIBERT checkpoint [1]. These embeddings are then passed to the decoder along with the expected output which in this case, will be the questions to be generated. The questions, like the paragraph, are passed through a tokenizer first and then given as input to the decoder as single question in the following form:

$$[\text{CLS}], q_1, q_2, \ldots, q_m, [\text{SEP}]$$

The questions are then generated as the output of the decoder. The questions and their corresponding context are then passed through a pre-trained Bert For Question Answering model which extracts the answer from the given context for the question (Fig. 1).

The obtained question and corresponding answer are then used to generate distractors, i.e. the alternate options given in multiple choice questions. For the purpose of distractor generation, the Word2Vec [8] corpus is used. The distractor generation was done using Google's pre-trained model of Word2Vec [8].

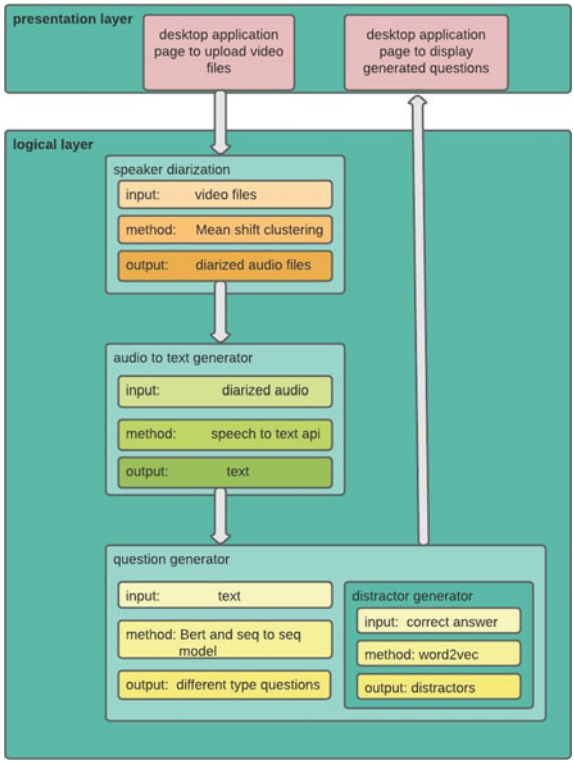


Fig. 1 System architecture

Some of the factors that pose a potential challenge include the external noise in the audio of the lecture, out-of-vocabulary words in the input to question generator, and handling various forms of answers for distractor generation.

6 Experimental Result

6.1 Experimental Setup

The speaker diarization mechanism consists of an external library that assists in clustering the data received from the audio/video file. By using mean shift clustering algorithm, we find the cluster with the most number of segments associated with it and assume it to be the one associated with the primary speaker's voice. The original file contents are then separated into separate smaller files containing the speech segments containing only speaker's dialogues. These segments are passed on to the Google Speech-to-text API to obtain their transcripts.

The question generation was implemented using the SCIBERT [1] uncased checkpoint to initialize both the encoder and decoder as a SCIBERT2SCIBERT network. The implementation is done using the Seq2Seq Transformer network from HuggingFace. The model is fine-tuned over the SQuAD training dataset [18], and then tested first over the SQuAD test dataset [18] followed by generation of questions for domain-specific contexts derived from video lectures. The parameters fine-tuned for the model include a maximum length of 30 for decoder generation, length penalty of 0.25, number of beams for Beam Search set to 5, and temperature parameter set to 0.95. The model is trained over the dataset for 3 Epochs, carrying out evaluation every 4000 steps on a GPU. The SCIBERT tokenizer is used to create the embeddings for the inputs to the encoder and decoder.

The distractor generation was done using Word2Vec [8] $most\_similar$ function with arguments as correct answer and top n set to 20 and out of the generated most similar words the least three similar words are taken as distractors. This ensures that the distractors are neither too similar nor too different with respect to correct answer.

6.2 Results

The speaker diarization and speech segmentation code generates multiple small audio segment files to be processed individually by the Google speech-to-text API. It is accurate to about 78%, based on analysis of 100 trials. The model successfully generates non-degenerative wh- questions from the given context. A few of the generated questions for the SQUAD test dataset [18] are shown in Table 1. From the examples, we can see that a majority of the questions, while contextually correct,

Table 1 Questions generated for input from SQuAD test data [18]

Snippet of context	Expected question	Generated question	Remarks
Some theories of civil disobedience … Brownlee argues … organizations and foreign governments …	What do some theories claim about civil disobedience?	What does Brown believe should be justified against?	The question contextually is correct but replaces the correct name "Brownlee" with the name "Brown"
In 1755, six colonial governors in North America met with General Edward Braddock, … New England …	When did colonial governors meet with General Edward Braddock about attack on the French?	Who was the British army commander of North America?	The question can be answered for the given context and grammatically correct
The English name "Normans" comes from … Latinized variously as Nortmannus, … to mean "Norseman, Viking"	What is the original meaning of the word Norman?	In medieval Latin, what is the English term for Normans?	Question is vague in terms of meaning. Required answer cannot be ascertained immediately

Table 2 Questions generated by model for domain-specific paragraphs

Input text	Generated question	Generated answer
If we define artificial intelligence is the ability … predict the future and … fields	What is the ability of a machine to predict the future?	Artificial intelligence
Hadoop is a famous big data framework	What is a famous big data framework?	Hadoop
Speaker diarization … includes a number of speakers into homogeneous segments	What is speaker diarization?	The technical process of splitting up an audio recording stream

deviate from the expected questions for the given paragraph context as per the dataset, this results in lower precision and BLEU score of 5.12 for the model over the SQUAD dataset [18].

The module is successfully able to generate questions for inputs given from the field of computer science as shown in Table 2. The model takes into consideration the context of the question even in the presence of abbreviations within the input. The contexts obtained from the transcripts of the input video were also successfully used to create questions, answers, and distractors. However, the model showed some difficulty in generating questions whenever the input consisted of any slang or colloquialisms of speech. For example, the phrase "sure-fire way" may be misinterpreted as a context about fire.

Table 3 Sample generated distractors

Question	Answer	Distractors
What is the ability of a machine to predict the future?	Artificial intelligence	1. Neural networks 2. Autonomous robots 3. Mathematical algorithms
How is hydrogen oxidated?	Liquid oxygen	1. Cryogenic engine 2. Supercold propellants 3. Propellants
When was the vax family introduced?	1978	1. 1993 2. 1992 3. 1991
What is a program called that enables a large scale across hundreds of thousands of millions of times?	MapReduce	1. Postgres 2. NoSQL 3. RDBMS

In case of Multiple choice questions the distractor generator is used which successfully generates distractors using the correct answer as input. A few generated distractors along with question and answer, which are generated on SQUAD test dataset [18] are shown in Table 3. Since we are using Word2Vec [8] for the distractor generation, the model cannot generate distractors for input having multiple words (more than two words). In such cases, only the question is given as output.

7 Conclusion

The system will allow the question generation model to take only the relevant speaker into consideration in videos containing multiple voices. Further, the usage of the mean shift algorithm allows the diarization to be carried out even if the number of speakers is not known. Another advantage of the methodology is that the usage of SCIBERT will allow terms specific to the scientific domain to be processed in a more accurate manner. Overall, the system can aid students in revising and assessing their knowledge while transitioning to a new mode of education. There are a few areas of the system that require further study such as the system's handling of inputs containing colloquial terms, in which case, the result may be inaccurate for the given context. Further, there is only one question generated per speech segment, which might lead to exclusion of any other concepts in the given segment, besides the one considered.

Future directions that can be explored for our system include a ranking system for the generated questions in terms of relevance to the entire transcript, a copy mechanism for handling domain-specific keywords, and expansion of input domains beyond science subjects.

References

1. Beltagy I, Lo K, Cohan A (2019) SciBERT: a pretrained language model for scientific text. arXiv preprint arXiv:1903.10676
2. Chan YH, Fan YC (2019) A recurrent BERT-based model for question generation. In: Proceedings of the 2nd workshop on machine reading for question answering, pp 154–162
3. Das B, Majumder M, Phadikar S, Sekh AA (2021) Automatic question generation and answer assessment: a survey. Res Pract Technol Enhanc Learn 16(1):1–15
4. Devlin J, Chang MW, Lee K, Toutanova K (2018) BERT: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805
5. Du X, Shao J, Cardie C (2017) Learning to ask: neural question generation for reading comprehension. arXiv preprint arXiv:1705.00106
6. Fujita Y, Kanda N, Horiguchi S, Xue Y, Nagamatsu K, Watanabe S (2019) End-to-end neural speaker diarization with self-attention. In: 2019 IEEE automatic speech recognition and understanding workshop (ASRU). IEEE, pp 296–303
7. Gish H, Siu MH, Rohlicek JR (1991) Segregation of speakers for speech recognition and speaker identification. In: ICASSP, vol 91, pp 873–876
8. Goldberg Y, Levy O (2014) word2vec explained: deriving Mikolov et al.'s negative-sampling word-embedding method. arXiv preprint arXiv:1402.3722
9. Gorodetski A, Dinstein I, Zigel Y (2019) Speaker diarization during noisy clinical diagnoses of autism. In: 2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, pp 2593–2596
10. Heilman M, Smith NA (2009) Question generation via overgenerating transformations and ranking. Tech. rep, Language Technologies Institute, Carnegie-Mellon University, Pittsburgh, PA
11. Huang Z, Watanabe S, Fujita Y, García P, Shao Y, Povey D, Khudanpur S (2020) Speaker diarization with region proposal network. In: ICASSP 2020–2020 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 6514–6518
12. Kenny P, Reynolds D, Castaldo F (2010) Diarization of telephone conversations using factor analysis. IEEE J Sel Top Signal Process 4(6):1059–1070
13. Liu Y (2019) Fine-tune BERT for extractive summarization. arXiv preprint arXiv:1903.10318
14. Ma L, Ma Y (2019) Automatic question generation based on MOOC video subtitles and knowledge graph. In: Proceedings of the 2019 7th international conference on information and education technology, pp 49–53
15. Mitkov R, Le An H, Karamanis N (2006) A computer-aided environment for generating multiple-choice test items. Nat Lang Eng 12(2):177–194
16. Qu C, Yang L, Qiu M, Croft WB, Zhang Y, Iyyer M (2019) Bert with history answer embedding for conversational question answering. In: Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval, pp 1133–1136
17. Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I et al (2019) Language models are unsupervised multitask learners. OpenAI Blog 1(8):9
18. Rajpurkar P, Zhang J, Lopyrev K, Liang P (2016) Squad: 100,000+ questions for machine comprehension of text. arXiv preprint arXiv:1606.05250
19. Rothe S, Narayan S, Severyn A (2020) Leveraging pre-trained checkpoints for sequence generation tasks. Trans Assoc Comput Linguist 8:264–280
20. Scialom T, Piwowarski B, Staiano J (2019) Self-attention architectures for answer-agnostic neural question generation. In: Proceedings of the 57th annual meeting of the association for computational linguistics, pp 6027–6032
21. Srinivas MJ, Roy MM, Kumar V (2019) Towards generating plausible distractors for code comprehension multiple-choice questions. In: 2019 IEEE tenth international conference on technology for education (T4E). IEEE, pp 19–22
22. Wang Z, Lan AS, Nie W, Waters AE, Grimaldi PJ, Baraniuk RG (2018) QG-net: a data-driven question generation model for educational content. In: Proceedings of the fifth annual ACM conference on learning at scale, pp 1–10

23. Wu X, Jiang N, Wu Y (2020) A question type driven and copy loss enhanced framework for answer-agnostic neural question generation. In: Proceedings of the fourth workshop on neural generation and translation, pp 69–78
24. Zavala L, Mendoza B (2018) On the use of semantic-based AIG to automatically generate programming exercises. In: Proceedings of the 49th ACM technical symposium on computer science education, pp 14–19
25. Zhou Q, Yang N, Wei F, Tan C, Bao H, Zhou M (2017) Neural question generation from text: a preliminary study. In: National CCF conference on natural language processing and Chinese computing. Springer, pp 662–671

Detecting Depression in Tweets Using Natural Language Processing and Deep Learning

Abhishek Kuber, Soham Kulthe, and Pranali Kosamkar

Abstract COVID-19 has caused physical, emotional, and psychological distress for people. Due to COVID-19 norms, people were restricted to their homes and could not interact with other people, due to which they turned to social media to express their state of mind. In this paper, we implemented a system using TensorFlow, which consists of multilayer perceptron (MLP), convolutional neural networks (CNN), and long short-term memory (LSTM), which works on preprocessing, semantic information on our manually extracted dataset using Twint scraper. The models were used for classifying tweets, based upon whether they indicate depressive behavior or not. We experimented for different optimizer algorithms and their related hyperparameters for all the models. The highest accuracy was achieved by MLP using sentence embeddings, which gave an accuracy of 94% over 50 epochs, closely followed by the other two.

Keywords Long short-term memory · Convolutional neural network · Multilayer perceptron · Depression detection · Social media analytics · Natural language processing

1 Introduction

Mental well-being is vital for the normal functioning of a human being. According to WHO [1], depression is distinguished from typical mood swings and relatively brief emotional responses to routine difficulties. Depression can be detrimental to one's health, especially if it persists and has a moderate or severe intensity. It can make the individual who is affected suffer severely and perform poorly at work and in school, etc. In extreme cases, depression can cause people to take extreme measures,

A. Kuber (✉) · S. Kulthe · P. Kosamkar
MIT World Peace University, Pune, India
e-mail: abhi.kuber@gmail.com

P. Kosamkar
e-mail: pranali.kosamkar@mitwpu.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_43

like suicides. Beyond 300 million individuals are affected due to depression, and just a portion of them obtain suitable therapy. Depression is one of the top causes of disability globally, with about 800,000 people dying by suicide each year, and is one of the main causes of death among persons aged 15–29 [2].

Amid the COVID-19 pandemic, many people were affected mentally due to the death of their loved ones, financial crisis, and social depression. Social interactions have been negligible, so people are moving toward social media to express their state of mind, while maintaining anonymity. In the World Happiness Report, India is ranked at 139 out of a total of 149 countries [3]. According to a survey conducted in India, there is a 70% rise in the people affected by depression and obsessive–compulsive disorder due to COVID-19. In India, depression is not classified as a disease due to the social stigma surrounding it, which prevents many affected individuals from seeking suitable professional assistance. By analyzing and setting some semantic markers in social media posts, it is possible to create a model that can give insight into a user's mental state, significantly sooner than current procedures, and can aid in the early detection of indicators of depression.

Early depression intervention programs, particularly for young individuals, can be transformed with the help of social media. Twitter is a social media platform where users can send and receive messages known as "tweets", and around 200 million of them are sent daily. Many people post their daily activities on social media, which is now an essential part of their daily lives [4]. Such real-time data on one's daily life could give priceless insights into one's emotions and behavior that structured surveys might not be able to provide [5]. The pilot study conducted by Park et al. [6] to assess the stability of a social networking site to be utilized as a source to identify depression suggested that 42.04% of tweets collected using the word "depression" conveyed one's negative emotions.

This paper examines a deep learning approach based on linguistic features which can be used to detect depression in Twitter users.

2 Related Work

Plenty of work is carried out in the domain of NLP for detecting depression from social media and adverse effects from post market surveys. The related work offers you an overview of the work that has been done in the field of depression analysis from text using various techniques.

AlSagri and Ykhlef [7] use machine learning techniques to detect probable depressed Twitter users based on his/her web activity and their tweets. The investigation of traits and their impact on detecting depression levels is the study's key contribution. Their results show that the more features they use, their models achieve higher scores. Support vector machine gives the best performance, with an accuracy of 82.5% and $F1$-score of 0.79. Islam et al. [8] focus on four types of factors, such as emotional process, temporal process, linguistic style, and all the three features together for detection of depressive text from Facebook posts. To investigate each

factor type, they used supervised machine learning techniques. The classification techniques used were decision tree, *k*-nearest neighbor, support vector machine, and ensemble, with decision tree giving the best performance, *F*1-score of 0.73. Singh et al. [9] construct neural-based architectures (RNN, GRU, CNN) to predict depression. They analyze performance of character-based versus word-based models and pretrained embeddings versus learned embeddings. The best performing models were the word-based GRU with an accuracy of 98% and word-based CNN, with an accuracy of 97%. Orabi et al. [10] try to find the best effective deep neural network architecture out of a few that are effective in NLP. CNN-based with max-pooling models provided better results than RNN-based models. The CNNWithMax models, using their optimized embedding, achieved a higher accuracy (87.957%), *F*1-score (86.967%), AUC (0.951), precision (87.435%), and recall (87.029%). Wen [11] applied a deep neural network that was applied to predict depression. The architecture consisted of an embedding layer followed by a linear layer. By optimizing the embedding dimension and learning rates, an accuracy of 99.45% was observed.

3 System Architecture

Because social media conversations take place in a more naturalistic atmosphere, it is crucial to figure out how much personal information an individual has shared and whether the information is true and sufficient.

3.1 Dataset

We work on tweets for indications of depressive behavior. We parse and export this to a JSON file, from which we build our input to the model. This was a very important and challenging task, as there are not any datasets available publicly. To tackle this challenge, we scrape Twitter until the desired number of tweets are reached. We utilize Twitter's lookup operators to scrape tweets on specific topics and hashtags and remove usernames from tweets so as to maintain anonymity. We utilized a collection of 204 depression-related phrases from a research published by Maigrot et al. [12]. From the list of 204 depression-related terms, we select a few trigger words and phrases, which could potentially indicate depressive behavior. These words in majority of the cases indicated depressive behavior from the tweet. The number of depressive terms in a tweet was calculated as a feature. All depression phrases are counted as a specific attribute rather than counting the frequencies of each term. The phrases are:

> alienated, antidepressant, better place without me, crying, death seeking, feeling blue, depressed, despair, done with life, fed up with life, hopelessness, im crying, kill myself, melancholia, self destructive, suicidal, suicidal thoughts, end it all, i feel worthless.

In addition to this, we select the language as English and the location as India. The scraper does not always give tweets in the English language, so we remove those tweets in the English language. On exploration of the dataset, we also found that some people tweet in local languages, using the English alphabet. These tweets were removed as well. Finally, we have around 17,000 tweets, t which we apply some preprocessing. We remove the mentions of a user (removing usernames), so as to maintain anonymity. Links are removed, as they do not really help in our task. Finally, all characters are converted to lowercase, and only alphanumeric characters are kept. In order to avoid class imbalance, random tweets on different topics were added as well.

3.2 Annotating the Dataset

The dataset was manually annotated, with the automatic annotation producing a lot of false positives. The simple approach would be to label a tweet containing any of the trigger words/phrases as 1 (indicating depressive behavior), but this method gives out a lot of false positives. Take an example from our dataset,

what is chelsea doing?? makes me want to kill myself

Although this contains the trigger phrase, the intention of this tweet is not to indicate depressive behavior. Hence, manual annotation was the best method.

Table 1 shows how the labels are distributed over the dataset. The tweets which indicated depressive behavior were labeled as 1, and the others were labeled as 0.

4 Methodology

For implementing the system, we have used LSTMs, MLPs, and CNNs for detecting depression through social media text (Fig. 1).

Table 1 Distribution of dataset

Dataset information	Number of examples
Total number of tweets	17,620
Tweets indicating depressive behavior	6774
Tweets indicating non-depressive behavior	10,846

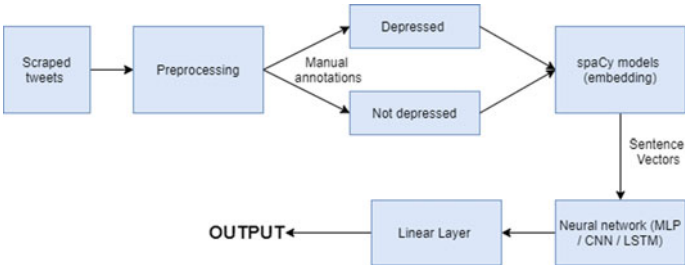


Fig. 1 Flow diagram of the system architecture

4.1 Long Short-Term Memory Networks

LSTM [13] is an "artificial recurrent neural network (RNN) architecture", which is skilled at learning long-term dependencies and works well on sequential data. While RNNs have difficulty with long-term dependencies due to the vanishing gradient, LSTMs are specifically created to prevent the long-term dependency problem. In our model, the tokenizer turns text into a series of numbers, each of which represents a word's index in the internal vocabulary. Although the numbers do not include any semantic information, the data are still sequential, therefore this works.

4.2 Convolutional Neural Networks

Instead of pixels, the text is represented as a matrix, which is the input to the model [14]. Each row in the matrix represents a sentence, as here, we are using sentence embeddings, but they could be words or even characters (for character level CNNs). The filters move across the entire matrix row by row. In images, we see that pixels close to each other are similar to each other. But, in the case of text, that is not the case. Looking at this problem, CNNs should not work for text related tasks, but it turns out CNNs work well. Text as sentence embeddings is passed to the CNN. Changing filter sizes are applied to each layer and followed by GlobalMaxPooling1D. The output is then flattened, which goes into a dense layer, dropout, and then to the output layer. Building a CNN means there are some hyperparameters to choose from. We investigate different parameters by holding runs over the dataset. We choose Adamax with a learning rate of 0.001 (check graphs in the results section).

4.3 Multilayer Perceptron

Word embeddings are a way of representing words in the form of real-valued vectors. Words that are similar to each other lie closer to one another mathematically in the

vector space. Cat and Dog would be closer to each other than Cat and Car. There are different algorithms, pretrained ones such as Word2Vec or GloVe [15]. There are also ways to learn embeddings, but for that, one would need a large amount of text, somewhere around millions of words. There are ways to work with sequences of text rather than just words. Sentence embeddings represent entire sequences and also capture semantic information. One of the simple ways to get sentence embeddings is to get all the word vectors of the words in a sentence and aggregate them together to get its respective sentence embedding.

5 Performance Analysis

The metrics used for evaluating our methods are accuracy and weighted averages of precision, recall, and $F1$-score.

- Calculated as the proportion of true positives to the summation of true positives and false positives for each class.

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \tag{1}$$

- Calculated as the proportion of true positives to the summation of true positives and false negatives for each class.

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \tag{2}$$

- The $F1$-score is calculated as below

$$F1\ \text{score} = 2 * \text{Recall} * \text{Precision}/(\text{Recall} + \text{Precision}) \tag{3}$$

5.1 Multilayer Perceptron

We first set the number of epochs to 50 and experimenting for different optimizer algorithms. Once we finalize the algorithm, we do the same for its learning rate and momentum. This method performs very well. It gives a precision, recall, $F1$-score, and an accuracy of 0.94 (Fig. 2).

5.2 Convolutional Neural Network

We use recurring units of convolution, max-pooling, and batch normalization. The number of epochs was set to 20, and different optimizers and their learning rates

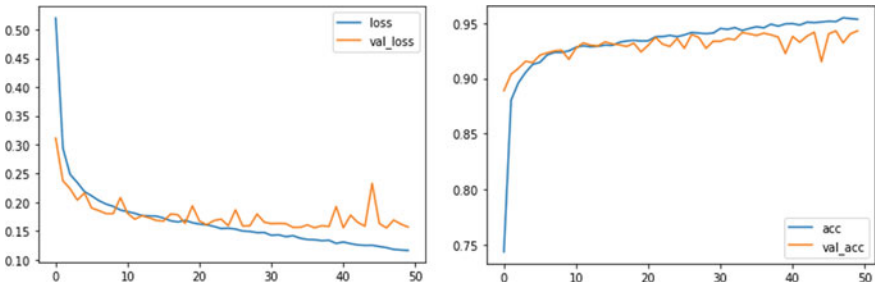


Fig. 2 Loss and accuracy plots for MLP, with SGD optimizer with a learning rate of 0.01 and a momentum of 0.4

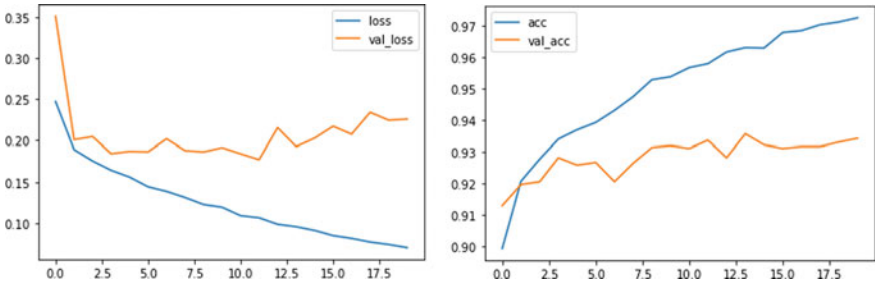


Fig. 3 Loss and accuracy plots for CNN, with Adamax optimizer and a learning rate of 0.001

were tried out, and Adamax optimizer with a learning rate of 0.001 gives the best results of precision, recall, $F1$-score, and accuracy of 0.93 (Fig. 3).

5.3 Long Short-Term Memory Network

LSTMs take longer time to train. Different optimizers and their learning rates were used over 5 epochs. It is observed that Adamax with a learning rate of 0.001 is the best fit and gives a precision, recall, $F1$-score, and accuracy of 0.93 (Fig. 4; Table 2).

6 Conclusion

In this paper, models for detecting depression in tweets using semantic information were implemented. Three architectures were used, MLP, CNN, and LSTM, which were fed sentence embeddings as input. The best performing model was the MLP, with an accuracy of 94%. The models mentioned in this paper can be used by mental health professionals for detection of depression on various social media platforms and

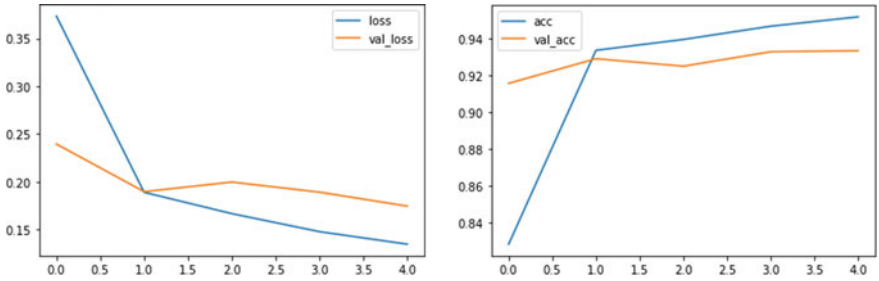


Fig. 4 Loss and accuracy plots for LSTM, with Adamax optimizer and learning rate of 0.001

Table 2 Algorithms and their results

Algorithm	Precision	Recall	*F*1-score	Accuracy
MLP	0.94	0.94	0.94	0.94
CNN	0.93	0.93	0.93	0.93
LSTM	0.93	0.93	0.93	0.93

for timely detection before the conditions get worse. This could potentially reduce healthcare costs and would not limit the concerned individual due to undiagnosed mental illnesses. In the future, we can classify the user tweets according to different levels of severity. This experiment could further mark the infrastructure for new procedures applied in different areas of health care and medical facilities to detect depression at an early stage for timely diagnosis.

References

1. World Health Organization. Depression. https://www.who.int/news-room/fact-sheets/detail/depression. Accessed 28 June 2021
2. You are what you tweet—towards data science, medium. https://towardsdatascience.com/you-are-what-you-tweet-7e23fb84f4ed. Accessed 27 June 2021
3. World happiness report 2021. https://worldhappiness.report/ed/2021/. Accessed 24 June 2021
4. 200 million tweets per day. https://blog.twitter.com/en_us/a/2011/200-million-tweets-per-day. Accessed 30 June 2021
5. Monitoring tweets for depression to detect at-risk users. https://ruor.uottawa.ca/bitstream/10393/36030/1/Jamil_Zunaira_2017_thesis.pdf. Accessed 29 June 2021
6. Depressive moods of users portrayed in Twitter. https://nyuscholars.nyu.edu/ws/portalfiles/portal/90894123/depressive_moods_kdd.pdf. Accessed 28 June 2021
7. AlSagri H, Ykhlef M (2020) Machine learning-based approach for depression detection in Twitter using content and activity features. IEICE Trans Inf Syst E103.D(8):1825–1832. https://doi.org/10.1587/transinf.2020EDP7023
8. Islam MR, Kabir MA, Ahmed A, Kamal ARM, Wang H, Ulhaq A (2018) Depression detection from social network data using machine learning techniques. Health Inf Sci Syst 6. https://doi.org/10.1007/2Fs13755-018-0046-0

9. Detecting depression through tweets. https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1184/reports/6879557.pdf. Accessed 28 June 2021
10. Orabi AH, Buddhitha P, Orabi MH, Inkpen D (2018) Deep learning for depression detection of Twitter users. Int J Comput Linguist Appl 88–97. https://doi.org/10.18653/v1/W18-0609
11. Wen S. Detecting depression from tweets with neural language processing. J Phys Conf Ser 1792. https://doi.org/10.1088/1742-6596/1792/1/012058
12. Maigrot C, Bringay S, Azé J (2016) Concept drift vs suicide: how one can help prevent the other? Int J Comput Linguist Appl
13. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
14. Kim Y (2014) Convolutional neural networks for sentence classification. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1746–1751. https://doi.org/10.3115/v1/D14-1181
15. Pennington J, Socher R, Manning C (2014) Global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543. https://doi.org/10.3115/v1/D14-1162

Predicting the Consumer Behaviour Based on Comparative Analysis of Salary and Expenditure Using Data Mining Technique

Teena Vats and Kavita Mittal

Abstract Data mining is also called information mining (otherwise called information revelation from datasets) and is the interaction of extraction of covered up, already obscure and possibly helpful data from datasets. The result of the separated information can be dissected for the future arranging and improvement points of view. In this paper, researcher has created an endeavour to exhibit the relation between monthly income and expenditures for calculating the customer behaviour. Their investigation utilizes the amazing information data mining tool Weka.

Keywords Consumer behaviour · Data mining · Visualization · Weka · Income · Expenditure

1 Introduction

To extricate concealed data from a huge dataset or an information stockroom is the lone way and is called information mining (DM). Information mining is approach to uncover the concealed data from the enormous information. Information mining has new and amazing innovation that help organizations and association to zero in on recover important data from their information distribution centre. Information mining is filling in as apparatus that assists with anticipating the future patterns and client conduct. Information mining permitting organizations to make proactive, information-driven choices, it has capacity to offer the response of that customary inquiry which is extremely tedious to determine. DM is discovering those examples which are missed by the specialists on the grounds that these examples are past the master's possibility. Information mining is an interaction of revelation design for information disclosure, it information from different sources. After the interaction

T. Vats (✉)
Guru Gobind Singh Indraprastha University, Dwarka, Delhi, India
e-mail: tina.summer27@gmail.com

K. Mittal
Jagannath University NCR, Bahadurgarh, Haryana, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_44

of assortment of the information, digging examination the information for future forecast. Means information digging utilize past information for future forecast. DM stores information in different measurements, classes and separates that information for building the example in various points and viewpoints. DM dissecting the example in enormous clusters so more than one programming utilize that example at same coordinated has different application in different fields, similar to clinical, entertainment, instruction, games, business, science, mechanical, research, etc. This study is about the large business use of the information mining. By utilizing information mining application scientist can more comprehend of their buyer and their purchasing conduct subsequent to learning the customer needs specialist ready to assemble more compelling example and strategy that impact purchaser for buying, it is conceivable simply because of information emulating. Information mining centre around ordinary shopper as well as spotlight on no standard customer organization for the most part use DM for monetary, retail, correspondence and advancement. Its centre around the relationship of item implies how they are related to one another, and it helps the plan of the thing in the stores that help the purchaser and valuable to the association. On the off chance that the scientist discovers the relationship of the thing in right manner, mastermind then in an appealing way or plan. Legitimate game plan of thing and staff expertise and replacement draw in the purchaser. For instance, legitimate game plan of staple thing like pickle, jam, ketchup in combo impacts costumer to purchase. Affiliation precludes help to track down this sort of example. It discovers relationship of things and relationship between thing and client [1, 2].

1.1 Meaning of Customer

Costumer is an individual who makes buys products or administrations which he wants from the market. The customer settles on a buy choice of the great or administration subsequent to going through number of choices present on the lookout. He chooses which great or administration best suits his necessities then, at that point goes for its buy. Purchasing conduct of the client is only dynamic interaction or the demonstrations that are engaged with purchasing and utilizing of the bought merchandise. Purchasing conduct of the customer is affected by numerous a variables, for example, organization advertising system in drawing in the clients impacts them so that the client goes for the acquisition of the great or administration of that organization. The more the quantity of purchasers of the item, the more will be the accomplishment of the firm who produce them. For an organization to fabricate the merchandise or administration that is wanted by the client, first the organization ought to examine the requirements of the client. Organization ought to attempt an overview to comprehend the necessities of the client and should make the merchandise which best fulfils such need. Client generally goes for those items which are accessible in the market by investigating the options accessible. It is significant that the advertisers attempt to comprehend of what is the need of the client, where there

is a more prominent need of that specific great or administration, when the great must be presented in the market and how the clients respond towards that specific great or item. For offering an item to a buyer, it is fundamental that an affiliation ought to be set-up between both the merchant and the shopper. For building such a connection, it is exceptionally fundamental that the dealer should be perceived the necessities of the client in ideal manner as to fulfil them which would profit the two of them. An inappropriate comprehension of the necessities of the purchaser would leave the organization in misfortune since no buyer liked to go for buying of such things or office. In such an express, the buyer would go for the item fabricated from some other organization to fulfil his need. The above portrayal obviously expresses that lone those organizations will support in the market which best fulfils the client needs [3].

1.2 Customer Behaviour

The investigation of shopper conduct incorporates the learning of persons, group, organization, and this measure gets familiar with the purchasing tendency for customer and, furthermore, gains proficiency with his/her needs, society, financial, condition, companion's, circle and chooses. Individual conduct impacts by bunch or encompassing environment, for example, an understudy impact by his/her companion's circle for picking their wearing garments. Each client act diversely at the hour of shopping analyst never say that all customer act same at season of buying of a similar item. Few are fulfilled, yet not many are as yet befuddled at the time charging. A customer is influenced by different elements, and it could be influenced by an individual, by society, by culture, by brain science and by cast. There are some more factors that influence the costumer at the hour of shopping. These are not many variables [4, 5].

- Need of customer
- Society of customer
- Desire of customer
- Financial state of customer
- Education of customer
- Age of customer
- Designation of customer [3, 5].

2 Related Work

1. Jagtap and Kodge [6] in this paper, authors describe that they created an endeavour to show how one can remove the nearby (locale)-level registration, financial- and populace-related different information for information disclosure

and their examination utilizing the incredible information mining tool Weka. This paper describes the education ratio according to gender (male and female).

2. Ahmeda and Shehaba [7] in this paper, the researcher examines the costumer conduct by utilizing different information mining characterization methods. Analyst investigation of the client conduct applies various information mining techniques on same information and checks which method is ideal. In this examination, a proposal framework is utilized with the assistance of this framework where a client effectively searches a thing. Every exchange contains the information of the client for additional exchange to anticipate the following buying. In this investigation Weka is utilized as an instrument for arrangement calculation 10 overlay approval strategy is utilized. As per this investigation, most elevated precision rate is 87.13% of choice table and time taken to fabricate the model is 0.24. Bunching has the most reduced precision rate which is 36.30%, and 0.05 is to assemble the model.
3. Maheshwari et al. [1] in this paper, the author utilize the FP-growth algorithm. Author applied this calculation on essential informational index that is gathered from wholesalers or retailers. Confidence and backing factors are considered that fabricate a solid affiliation rule. FP-tree is produced as pre-handling task. Thing is arranged as a sliding request of recurrence, and less successive things are taken out. The after-effect of this investigation says what things are dealt together.
4. Ghogare and Kalmegh [8] as per the authors, this investigation reasons that Weka apparatuses are mining instruments. Weka is the assortment of different order calculations. Practically, same work is done by different mining apparatuses like moreover Orange, Rapid Miner, Weka, jHepWork, KNIME. Each device has its own upsides and downsides. Yet, all among Weka have its own advantage and amazingly important. It empowers a client to use Weka without any problem.
5. Pahwa et al. [2] as per the authors, this examination reasons that originations basically utilized data mining in a large scale of retail, finance showcasing associations and communications. Information mining has a tonne of significance as a result of its colossal materialness. Information mining assists the online retailers with arranging successful advertising methodologies, which will prompt development in the online market that will create benefits and further will prompt increment of the Indian economy.
6. Ijaz and Syafrudin [9] in this paper, the creator foresees client conduct based on constant information and affiliation rule for computerized signage-based online store (DSOS). According to this paper, different factors impact client conduct like age monetary status, education, environment. Researcher gathers information from different fields in past time like poll, interview, surveys, case study. In this paper, real-time information preparing and affiliation rule are utilized for investigation client conduct. MYSQL is used to store the backend information base. A client looks through the thing by utilizing the digital signage-based online store (DSOS). Examination of the client conduct is assisting the administration with taking better choice. The innovation utilized in huge information

for continuous information handling is NoSQL. MongoDB and Apache Kafka were used.

7. Gull and Pervaiz [5] in this paper, creator examines client conduct by their time spent on internet shopping site. The analyst analyses how many individuals are just riding the webpage or the number of individuals who buy the item and the number of individuals who add the thing on truck only. This task is finished by utilizing information mining methods and calculations. Apriori calculation utilized for the investigation and affiliation rule used to extricate the connection between the thing and fine out the number of product buy together. By this examination, analyst discovers the factor and backing factor of the things.

3 Data Mining

Data mining has been categorized like the nontrivial taking out of understood, already obscure and conceivably helpful data from datasets/information distribution centres. It utilizes AI, factual and representation methods to find and present information in a structure, which is effectively thorough to people. Information mining, the taking out of hidden perceptive data from huge datasets, is an incredible novel originality with unexpected possible to help customer centre just about the main data in their in sequence stockrooms. Information mining apparatuses predict future patterns and practices which authorize organizations to create practical, information-focused choices. The robotized, looming examinations obtainable by information mining shift precedent the investigations of previous occasions given by review strategy usual of choice psychologically supportive networks. Information mining devices can respond to business questions that generally were extremely monotonous to determine. They scour datasets for covered up designs, discovering prescient data "that high-quality may miss since it libel exterior their hypothesis. Information mining strategies can be accepted out rapidly on existing programming and tools stages to get better the benefit of alive data property, and can be integrated with new objects and frameworks as they are welcomed on-line" [2, 7, 10].

3.1 Mining Process

1. Information gathering: data is gathered from different information sources.
2. Information choice: select suitable information from gathered information which is valuable in mining.
3. Information tidiness: the chosen information might contain mistakes, missing qualities, uproarious or conflicting information. Thus, by applying various methods, these peculiarities will be cleaned.
4. Information alteration: after cleaning, information might require other change like smoothing, collection, standardization.

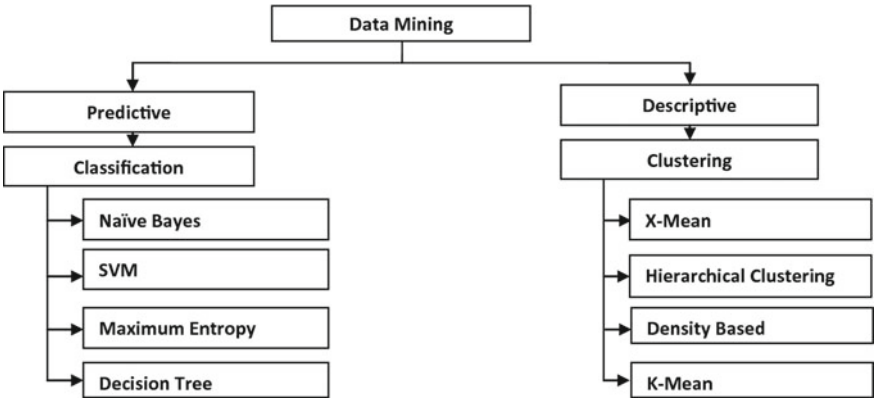


Fig. 1 Data mining technique

5. Information mining: now, information is prepared to mine. In this way, apply various information mining procedures to find helpful examples from information.
6. Example evaluation and knowledge presentation: after discovering examples, representation and eliminating excess examples will be done, from produced designs.
7. Dynamic: use found information, to take significant choices. Information Mining Techniques: there are some significant information mining methods like association, classification, clustering, prediction, sequential patterns, decision tree, which are utilized to discover significant information from the information (Fig. 1).

Predictive data mining: this procedure is utilized to investigate the previous information and through that make expectation about what is to come. For instance, as indicated by item deal makes expectation about the benefit. Incorporates an assortment of factual like demonstrating, AI, that examination current and verifiable realities to make expectation about future occasions. It gives answers to the inquiries, for example, "what can occur?" in future. It incorporates different arrangement calculations like fake neural organization, choice trees, AI, fluffy rationales and so on that can be utilized for the forecast of future business or patterns in information and help in dynamic [11].

Classification: this method is utilized to group the information into predefined classes or portions. For instance, characterize the items as per its qualities.

Descriptive data mining: incorporate every one of the information conglomeration strategies and procedures that give knowledge into past information. It breaks down every past datum and gives answers to the inquiries, for example, "what has occurred till information".

Clustering: this procedure is utilized to make precious bunches of in sequence which have same qualities. For instance, organize the things of a alike group in one frame. Here, group arrangement utilizing the principles of customer age and article contract.

Association: this process is utilized towards discover the example in regards to the connection between things in a similar exchange. For instance, discover example of items that customer much of the time buy together [8, 11].

All this analysis will be useful in

- Market basket analysis in which the data procures is, "Which things the customer buys together?" This data will be valuable for chipping away at stock in the store, organizing store design and propelling things.
- Sales gauging which will review time-based models, that is, "If the customer buys a thing today, when he/she buy the thing again?" This will help the retailer with staying aware of suitable stock.
- Information base marketing is in which retailers keep up with client profile about customer conduct. It signifies "Regardless of whether the shopper purchase marked items or go to the occasional deal?" This information will be utilized for practical item advancement.
- Retail items Planning and Allocation will inspect the item-purchasing behaviours as indicated by its situation into store. As per that significance "Is the buyer purchase the specific item, since he/she purchases the local item?" This will assist the retailer with planning the format of the store.

4 Weka

Waikato environment for knowledge analysis (Weka) is a standard set-up of AI programming written in Java, made at the University of Waikato, New Zealand. Weka is free programming open under the GNU general public license. The Weka workbench contains an arrangement of portrayal gadgets and estimations for data examination and farsighted illustrating, alongside graphical UIs for basic induction to this value [12, 13].

Weka is an arrangement of AI estimations for handling certifiable data mining issues. It is written in Java and runs on essentially any stage. The computations can be either applied directly to a dataset or called from your own Java code.

The primary non-Java transformation of Weka was a TCL/TK front-finish to (generally outsider) demonstrating calculations executed in other programming dialects, in addition to information pre-processing utilities in C and a build file-based structure for operation AI tests. That exclusive form was essentially designed as a gadget for exploratory information from horticultural places; though the afterwards completely Java-based version (Weka 3), for which progression begin in 1997, is at present utilized in different application regions, exclusively for educational purposes and assessment. Benefits of Weka include:

I. Open user-friendliness below the GNU general public license.

II. Handiness, as it is entirely accepted away in the Java programming language and for that cause runs on basically a few cutting edges figuring stage.
III. An exhaustive assortment of information pre-processing and demonstrating methods.
IV. Usability because of its graphical UIs.

Weka upholds a few standard information mining assignments, all the more explicitly, information pre-processing, clustering, classification, regression, visualization and attribute variety. The whole of Weka's strategy is predicated through the considerate so as to the in order is available as a lonely level text or link, where all information tip is depict by a permanent number of merits (frequently, numeric or superficial quality, still some other attribute types are likewise upheld). Weka gives entry to SQL in sequence bases utilizing Java database connectivity and can contract through the conclusion returned through a dataset question. It is not fit for multi-social information mining, yet there is discrete programming for changing over an assortment of connected dataset tables into a solitary table that is reasonable for preparing utilizing Weka. Another significant region that is at present not covered by the calculations remembered for the Weka conveyance is grouping displaying [3, 12–14].

4.1 Limitation of Weka

1. Weka can handle only small datasets.
2. Weka not support some new algorithms.
3. Filter section absences few specific transformations tool.
4. If one value is missing in dataset, then user not able to use Weka algorithms.

5 Data Description

Data in Fig. 3 shows the data that is used for analysis in Weka. That data is taken from the Kaggle repository. This data contains five attribute customer id, gender, age, monthly income and spending score. This data contains records of 200 customers who purchased goods from November 2019 to May 2020. Kaggle provides data in various formats like CSV and ARFF. Research finds that data in CSV format, then changes it in ARFF format for analysis of that data and takes out the results [15] (Table 1).

Table 1 Data description

Data	Instances	Attribute	Data type	Data format	Attribute name
Customer data	200	5	Numeric	CSV	Customer id, gender, age, monthly income, spending score

6 Data Processing, Methodology and Outputs

Additionally, the information is altered over to attribute relation file format (ARFF) association to gauge in Weka. ARFF evidence is an ASCII text file that depicts a rundown of happening contribution a group of merits. ARFF files were formed by the machine learning project at the section of Computer Science of The University of Waikato for use with the Weka AI programming. This library describes the version of ARFF utilized with Weka variant 3.2 to 3.9.2; this is an increase of the ARFF plan as depicted in the in order mining book [11, 14].

In the wake of handling the ARFF record in Weka the rundown, everything being equal, measurements and different boundaries can be used as displayed in Fig. 2.

After processing the ARFF file in Weka, the list of all attributes, statistics and other parameters can be utilized as shown in Fig. 3. In the shown file, there are 200 customer data processed with different attributes (05) like customer id, gender, age, monthly income (k) and spending score (1–100) generated by mall customer data. The processed data in Weka can be analysed using different data mining techniques like classification, clustering, association rule mining and visualization algorithms. Figure 4 shows the few processed attributes which are visualized into a two-dimensional graphical representation. The handled information in Weka can be investigated utilizing various information mining strategies like classification, clustering, association rule mining, visualization.

Figure 4 shows visualization of all attribute of the mall customer data by processed mining Weka in. The data can be extricated regarding at least two acquainted connection of informational collection. In this cycle, researcher has made an effort to imagine the effect of income on monthly expenditure.

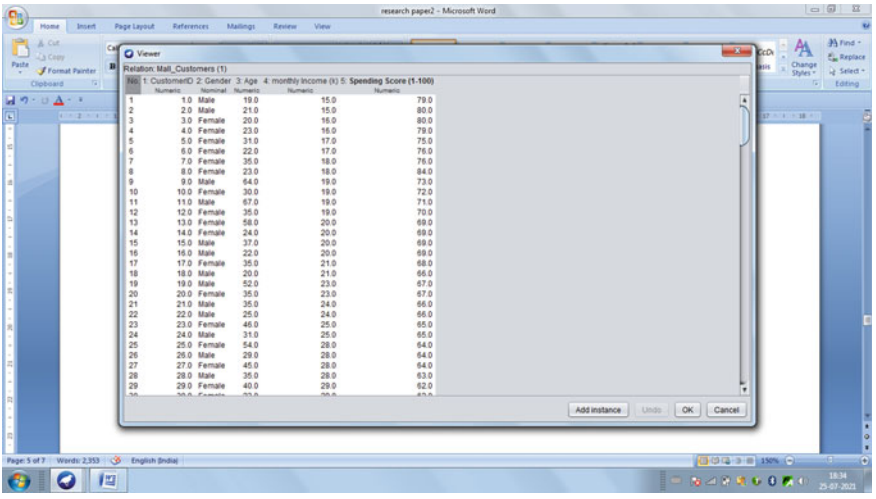


Fig. 2 Customer data in Weka

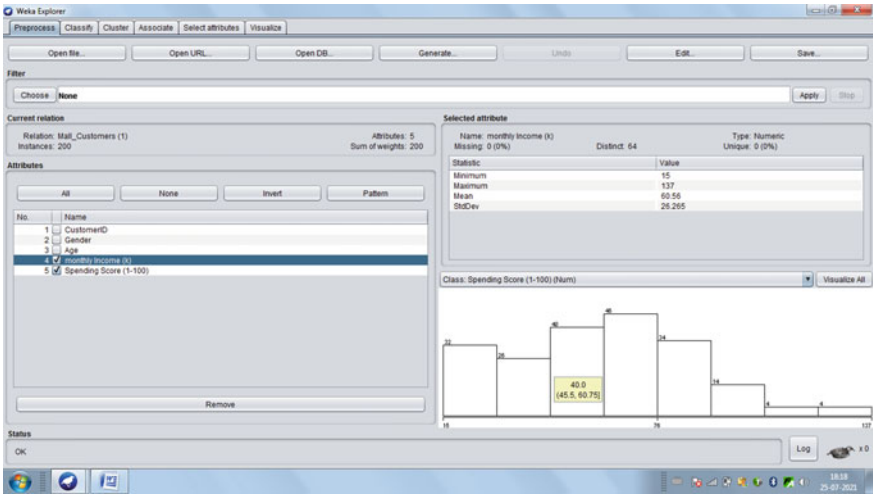


Fig. 3 Processed ARFF file in Weka

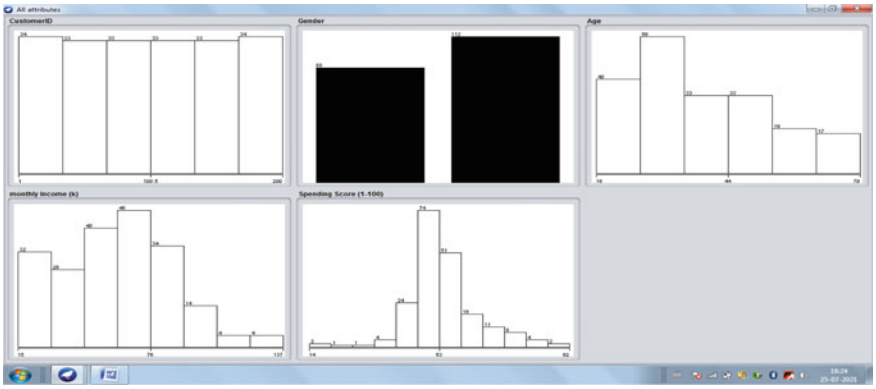


Fig. 4 Graphical visualization of all attributes

According to the data, result says that when the income becomes high, then the expenditure ratio is low approximately. Figure 5 shows that result. *X*-axis shows the expenditure per month by customer, and *Y*-axis shows the income of a customer per month.

7 Conclusion

Information extraction from dataset got one of the critical interactions of every single association for their advancement issues. This is not just significant in the business

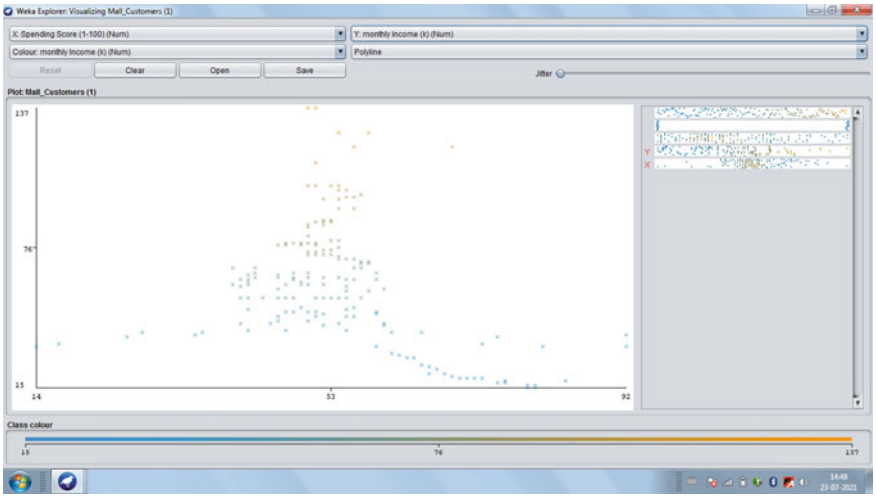


Fig. 5 Ratios of income and expenditure

ventures yet in addition assumes an essential part in e-commerce for expectations arranging and sequence matter. This study explains one of the slight significances of Weka to custom and analyse for listing in sequence mining matter as well as information revelation. It is for e-business. This study is helpful for managing the market according to the income because income decides the expenditure of the customer. It helps to decide the expenditure ratio according to income. E-commerce is the superset of the e-businesses. Ecommerce is further divided into two phases: one is e-business, and the other is e-services.

References

1. Maheshwari N, Pandey NK, Agarwal P (2016) Market basket analysis using association rule learning. Int J Comput Appl 975:8887
2. Pahwa B, Taruna S, Kasliwal N (2017) Role of data mining in analyzing consumer's online buying behaviour. Int J Bus Manag Invent 6(11):45–51
3. Tallapragada VS, Rao NA, Kanapala S (2017) EMOMETRIC: an IOT integrated big data analytic system for real time retail customer's emotion tracking and analysis. Int J Comput Intell Res 13(5):673–695
4. Chinchilla LDCC, Ferreira KAR (2016) Analysis of the behaviour of customers in the social networks using data mining techniques. In: 2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), Aug 2016. IEEE, pp 623–625
5. Gull M, Pervaiz A (2018) Customer behaviour analysis towards online shopping using data mining. In: 2018 5th international multi-topic ICT conference (IMTIC), Apr 2018. IEEE, pp 1–5
6. Jagtap SB (2013) Census data mining and data analysis using WEKA. arXiv preprint arXiv:1310.4647

7. Ahmeda RAED, Shehaba ME, Morsya S, Mekawiea N (2015) Performance study of classification algorithms for consumer online shopping attitudes and behaviour using data mining. In: 2015 fifth international conference on communication systems and network technologies, Apr 2015. IEEE, pp 1344–1349
8. Ghogare SA, Kalmegh SR (2019) Performance comparison of random forest and Hoeffding tree classifier using WEKA data mining tool on car reviews data. J Eng Res Appl 9(3):37–42. ISSN: 2248-9622
9. Alfian G, Ijaz MF, Syafrudin M, Syaekhoni MA, Fitriyani NL, Rhee J (2019) Customer behaviour analysis using real-time data processing: a case study of digital signage-based online stores. Asia Pac J Mark Logist
10. Raymond M, Francisco C, Fabio C, Jenny C (2020) A product-centric data mining algorithm for targeted promotions. J Retail Consum Serv 54:101940
11. http://en.wikipedia.org/wiki/Data_mining
12. http://www.cs.waikato.ac.nz/ml/weka/
13. http://www.cs.waikato.ac.nz/ml/weka/arff.html
14. http://en.wikipedia.org/wiki/Weka_(machine_learning)
15. https://www.kaggle.com/c/scl-2021-ds

Improve the Detection of Retinopathy with Roberts Cross Edge Detection

Arun Kumar Jhapate, Ruchi Dronawat, Minal Saxena, and Rupali Chourey

Abstract Image processing is most useful tool for pattern recognition and compare between two images. In retinopathy diagnosis, Roberts cross is an image processing filter that can be extracted vertical and horizontal edges, where the change has been detected from the corresponding pixels. It is a gradient kernel where vertical edges and horizontal edges are extracted separately, and later, magnitude has been computed by combining both the parameters. It has been considered as best edge detection tool in image processing for obtaining sensitive edges. For calculations, 150 images are taken for result which is generated by this system. Here, system achieved 96.66% of accuracy with nominal false acceptance rate. The results are run on MATLAB tool.

Keywords Diabetic retinopathy diagnosis · Image processing · Bioinformatics · Optic cup · Robert's edge detection · Optic disc · Retinal image detection

1 Introduction

Diabetic retinopathy means the disease that effect in eyes by diabetic present in blood. It creates light touches tissues insight of the eyes means retina brought about by harm to the veins of the light-touchy tissue at the rear of the eye (retina). In initial stage, diabetic retinopathy did not show any symptoms, but patient has only minor vision problems as time travels the symptoms unclear, instable, dark, empty, or loss of vision may develop which is very risky for any diabetic person. If a person has diabetic, then he has to regular checkup of eyes because if retinopathy is not diagnosis in early stage, then it can reach to blindness.

In medical science, there are two types of diabetic retinopathy, first is early diabetic retinopathy and second is advanced diabetic retinopathy. In early diabetic retinopathy

A. K. Jhapate (✉) · R. Dronawat · R. Chourey
Department of CSE, SIRT, Bhopal, MP, India
e-mail: arunjhapate11@gmail.com

M. Saxena
Department of EC, SIRT, Bhopal, MP, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_45

which is also represented by nonproliferative diabetic retinopathy (NPDR), new blood vessels are not growing and in advance diabetic retinopathy harmed veins chocked which causing the development of new, strange veins in the retina that is most dangerous for vision of eye. Managing blood sugar, diabetics, blood pressure, cholesterol, smoking, and pay attention to vision changes are some preventions for retinopathy, but we cannot always prevent diabetic retinopathy. Prevention is better than cure is the most important here.

In Fig. 1, it is clearly seeing that normal retina and diabetic retinopathy images are change. Here, image processing is used for diagnose the retinopathy. Several methods are available for diagnose the problem in retinopathy. The Roberts cross formula is used for edged detection, and he proposed a formula as human thinking psychologically and generate a method which works like a human edge detection. It was one of the primary edge markers and was at first given in 1963, by Roberts.

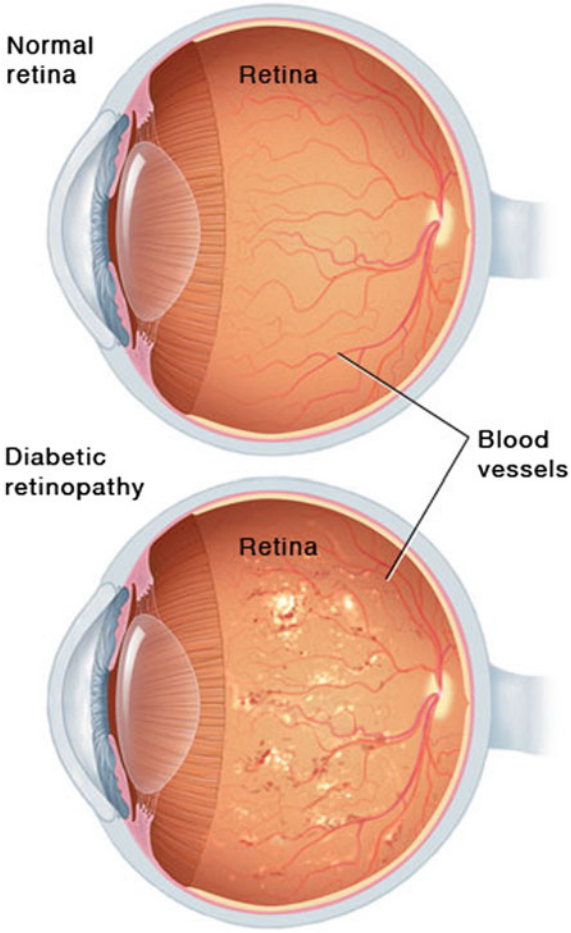


Fig. 1 Impairments over retinal image [1]

Roberts cross is an image processing filter that can be extracted vertical and horizontal edges, where the change has been detected from the corresponding pixels. It is a gradient kernel where vertical edges and horizontal edges are extracted separately, and later, magnitude has been computed by combining both the parameters. It has been considered as best edge detection tool in image processing for obtaining sensitive edges.

Roberts given the following equations for calculations:

$$y_{i,j} = \sqrt{x_{i,j}}$$
$$z_{i,j} = \sqrt{\left(y_{i,j} - y_{i+1,j+1}\right)^2 + \left(y_{i+1,j} - y_{i,j+1}\right)^2}$$

where x represents the value of initial intensity in the image,

z represents calculated derivative and i, j represents current pixel location in the image.

Above formula calculates the computed derivation in respect to value of i and j which represents the coordinates of pixel.

2 Related Works

Palavalasa et al. [2] proposed a procedure they used background and foreground subtraction method for subtraction they used fundus data set for apply their method. Procedure to select global threshold which is deepened on the global threshold value is selected by conducting several experiments such that the candidates include conditions present in the fundus images they used 89 images in their experiments and obtain 87% accuracy rate.

Ravishankar et al. [3] use an image from fundus data set and after morphological filling, they direct mark with "+", the main blood vessel and generate the final result after removing fovea, they got 84% accuracy results by using 10% patients having the retinopathy.

Gardner et al. [4] used healthy retinal image generated by real data set which available in DRIVE data set and improved the images by CLAHE tool then also took images from the DIARETDB0 data set apply same tool and generate enhanced images of data set. They used both data set for their evaluation and generate the results which generates the 89% accuracy result, they first time apply AI technique to solve the problem. It has been finished by utilizing 100 pictures of 47 patients with DR.

Alireza et al. [5] given a method which works on partition based on the color representation and used K-means algorithm for classification to their results, they used fuse c medium to generate the result, they generate 85.6% accuracy result for their result, and they segment the given problem in parts and apply the K-means so resultant value is changed and got more accuracy.

3 Problem Identification

Nowadays, visual deficiency is very common problem found in individuals. Detection of visual deficiency is done by subtracting background and foreground for covering the undesirable background. For doing this segment, the portion of vein by comparing threshold with entropy and altering over it into binary images. Yet, 0 and 1 chance that supply with or veiling has been end on middle age ranges between twenty's to sixty's years.

To detect diabetic retinopathy, the best way is consistently testing the fundus so this eye disease can be timely cure. In ophthalmology, various digitally programmed fundus cameras used to generate digital information which is useful for analyze the eye disease. Eye disease causes partial or complete loss of vision. Diabetic retinopathy is one of them. Medical experts can review these images and identify patients.

Here, a auto-generated system is proposed which detects the retinopathy detection in images and gives the result on the basis of images that it harm to eyes or not.

It highlights, the affirmation of retina images using legitimate picture preparing and once the shading modification has been done, then at that point, framework extracts the outlines by using Robert's cross edge detection and gives the results by comparing the throughput of images with threshold. When the edges have separated the n impedances will the results is shown.

4 Proposed Work

Here, we are developing a system where most popular method Robert's cross is used here; we are using horizontal and vertical kernel is used for edge detection the following kernel are used.

The most common kernels are in Figs. 2 and 3.

For following formula, it has been denoted as G_x and calculation can be performed as

$B11$ is multiply with horizontal coordinates values with metrics value which are $a11$, $a12$, $a21$, and $a22$; similarly for G_y calculations, all the vertical coordinators multiply with same metrics values and after calculations of final gradient magnitude will calculate with following formula.

Fig. 2 Kernel horizontal

+1	0
0	-1

Fig. 3 Kernel vertical

0	1
-1	0

$$G = \sqrt{G_x^2 + G_y^2}$$

The algorithm is as follows

Assumption: G_x and G_y are horizontal and vertical kernel, where x, y image pixel coordinators.

Here, we are taking fundus image as 2D image matrix as input and get gradient magnitude as output for applying; first convert the image into gray color, then apply color histogram on the image. After applying histogram and Robert's cross kernel, calculated throughput and compare that throughput to threshold if throughput is greater than threshold, then diabetic retinopathy detected otherwise treat as a normal image.

Here, I flowchart we ae taking total 150 retinal images.

From the data set and converted them in to gray color image which are the combination of 0–255 black and white combination, then apply Robert class kernel which works as horizontal and vertical and sum the solution for getting best result; after that, we calculated the throughput for image here if throughput is greater than the threshold, then the image is retinopathy otherwise it is normal image and algorithm end (Fig. 4).

Automatic diabetic retinopathy system is implemented in MATLAB and its taken input of fundus images (150 images) and then apply the algorithm the result are shown in Fig. 5.

5 Result Analysis

Here, the results are shown by Table 1, where data set is taken from fundus data set, and comparison result is shown. Taking 150 fundus images for calculating result that retinal infection is identified or not. The result generated 66 true positive (TP), 5 true negative (TN), 75 false positive (FP), and 4 false negative (FN) which is better than the previous results (Fig. 6; Table 2).

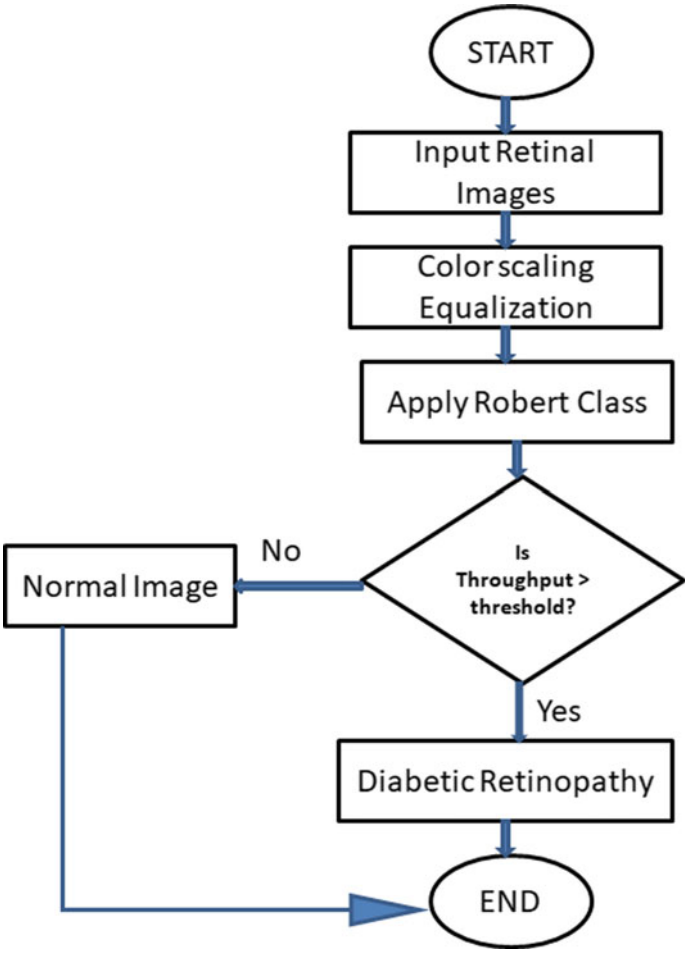


Fig. 4 Flowchart

6 Conclusion and Future Scope

This model gives the better solution we have conducted our experiment on MATLAB which is easy tool to use here we are taking 150 fundus images for calculating result that retinal infection is identified or not. The result generated 66 true positive (TP), 5 true negative (TN), 75 false positive (FP), and 4 false negative (FN) which is better than the previous results. The accuracy of system which was 87% by Palavalasa [2] and 95.5% by Sharma et al. [6]; here, our result is 96.66% where we are using throughput as comparison. The comparison result is represented by the result chart.

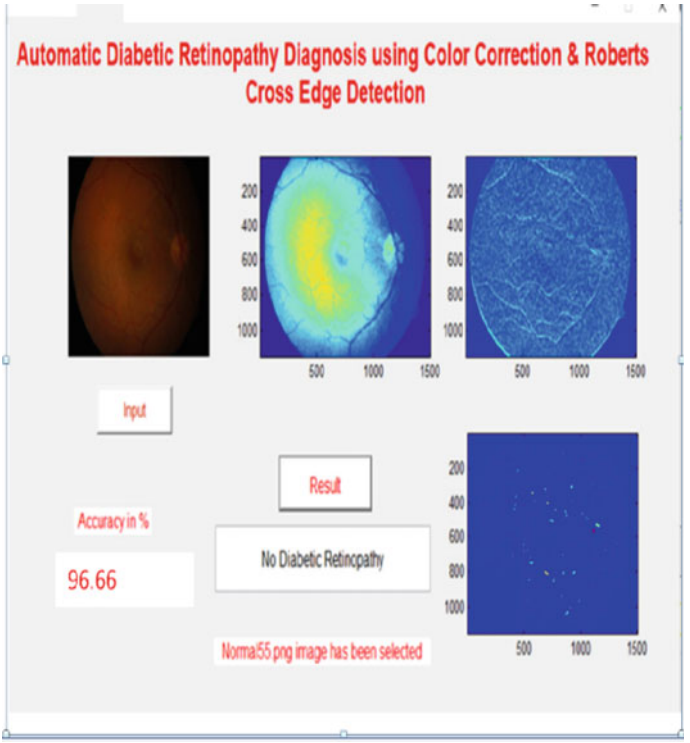


Fig. 5 Proposed work GUI

Table 1 Result comparison

	Palavalasa [2]	Sharma et al. [6]	Proposed algorithm
Data set	Fundus	Fundus	Fundus
Applied method	Subtracting background and foreground	Robert's cross edge detection with entropy comparison	Robert's cross edge detection with throughput comparison
Number of images	89	110	150
Result (accuracy) (%)	87.00	95.5	96.66

In future, we can take both the features throughput and entropy for best result. For better result, the number of images may high so we can check the result for large images and accuracy may be more.

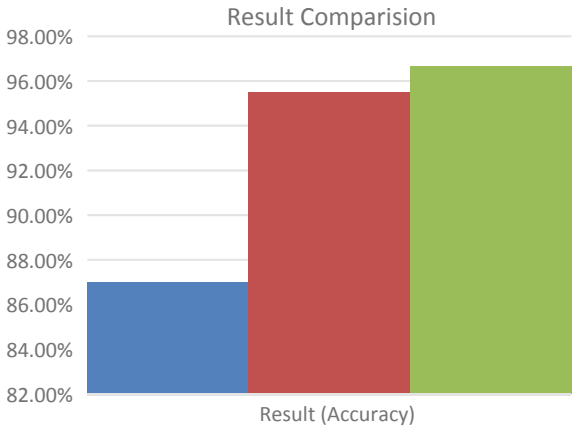


Fig. 6 Result chart

Table 2 Result and outcomes

	Outcomes	Result (accuracy) (%)
TP	66	96.36
TN	5	0.54
FP	75	0.03
FN	4	96.36
Total testing images	150	96.66

References

1. Ocutech. Diabetic retinopathy. https://ocutech.com/bioptics-help-individuals-with-visual-impairments/diabetic-retinopathy/. Accessed 18 Jan 2021
2. Chudler. Diabetic retinopathy. https://faculty.washington.edu/chudler/diabr.html. Accessed 17 Jan 2021
3. Argade KS, Deshmukh KA, Narkhede MM, Sonawane NN, Jore S (2015) Automatic detection of diabetic retinopathy using image processing and data mining techniques. In: 2015 international conference on green computing and internet of things (ICGCIoT), Noida, pp 517–521
4. Ravishankar S, Jain A, Mittal A (2009) Automated feature extraction for early detection of diabetic retinopathy in fundus images. In: 2009 IEEE conference on computer vision and pattern recognition, Miami, FL, pp 210–217
5. Gardner GG, Keating D, Williamson TH, Elliott AT (1996) Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. Br J Ophthalmol 80(11):940–944. https://doi.org/10.1136/bjo.80.11.940
6. Sharma A, Dronawat R, Jhapate AK (2021) Automatic diabetic retinopathy detection using Roberts cross edge detection in DIP. In: 2021 10th IEEE international conference on communication systems and network technologies (CSNT). ISSN: 2329-7182. https://doi.org/10.1109/CSNT51715.2021.9509675
7. Palavalasa KK, Sambaturu B (2018) Automatic diabetic retinopathy detection using digital image processing. In: 2018 international conference on communication and signal processing (ICCSP), Chennai, pp 0072–0076. https://doi.org/10.1109/ICCSP.2018.8524234

8. Sisodia DS, Nair S, Khobragade P (2017) Diabetic retinal fundus images: preprocessing and feature extraction for early detection of diabetic retinopathy. Biomed Pharmacol J
9. Osareh A, Shadgar B, Markham R (2009) A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images. IEEE Trans Inf Technol Biomed 13(4):535–545
10. Mukherjee A, Rathore D, Shree S, Jameel S (2015) Diagnosis of diabetic retinopathy. Int J Eng Res Appl 5(2), (Part-4). ISSN: 2248-9622
11. Trisha, Ali I (2020) Intensity based optic disk detection for automatic diabetic retinopathy. In: 2020 international conference for emerging technology (INCET), Belgaum, pp 1–5. https://doi.org/10.1109/INCET49848.2020.9154021
12. Solanki MS, Mukherjee A. Diabetic retinopathy detection using eye image
13. Khan MW (2013) Diabetic retinopathy detection using image processing: a survey. Int J Emerg Technol Res 1(1)
14. https://docs.opencv.org/3.4/Background_Subtraction_Tutorial_Scheme.png
15. Klein R, Klein BE, Moss SE, Davis MD, De Mets DL (1984) The Wisconsin epidemiologic study of diabetic retinopathy. II. Prevalence and risk of diabetic retinopathy when age at diagnosis is less than 30 years. Arch Ophthalmol 102:527–532
16. Harangi B, Lazar I, Hajdu A (2012) Automatic exudate detection using active contour model and region wise classification. In: IEEEE MBS 2012, pp 5951–5954
17. Harangi B, Antal B, Hajdu A (2012) Automatic exudate detection with improved naive-Bayes classifier, computer based medical systems. In: CBMS 2012, pp 1–4
18. Zuiderveld K (1994) Contrast limited adaptive histogram equalization. In: Graphics gems IV. Academic Press, pp 474–485
19. Langroudi MN, Sadjedi H (2010) A new method for automatic detection and diagnosis of retinopathy diseases in color fundus images based on morphology. In: International conference on bioinformatics and biomedical technology, pp 134–138
20. Giribabu KB, Venkata Subbaiah P, Satya Savithri T (2008) Segmentation of exudates and optic disk in retinal images. In: Sixth Indian conference on computer vision, graphics image processing

A Study of Emerging IoT Technologies for Handling Pandemic Challenges

M. Deepika and D. Karthika

Abstract Modern years, the Internet of Things (IoT) is mechanizing in abundant real-world functions such as smart transportation, smart business to build an individual life more accessible. IoT is the mainly used method in the previous decade in different functions. Deadly diseases always had severe effects unless they were well controlled. The latest knowledge with COVID-19 explains that by using a neat and speedy approach to deal with deadly diseases, avoid devastating of healthcare structures, and reduce the loss of valuable life. The elegant things are associated with wireless or wired communication, processing, computing, and monitoring dissimilar real-time situations. These things are varied and have low remembrance, less processing control. This article explains a summary of the system and the field of its function. The recent technology has supplied to manage previous closest diseases. From years ago, scientists, investigators, physicians, and healthcare specialists are using novel computer methods to resolve the mysteries of disease. The major objective is to study dissimilar innovation-based methods and methods that support handling deadly disease challenges that are further appropriate developments that can probably be utilized.

Keywords Artificial intelligence (AI) · Big data · Computational technologies · Internet of Things (IoT) · Pandemic diseases · COVID-19

1 Introduction

The coronavirus infection 2019 (COVID-19) from Wuhan, China, had been represented in December 2019 [1]. The illness is spreading bit by bit, and the social protection structure fights to manage each polluted individual, particularly in an

M. Deepika (✉)
Department of Computer Science, Crescent University, Chennai, India
e-mail: deepika.ca@crescent.education

D. Karthika
Department of Computer Science, SDNB Vaishnav College for Women, Chennai, India
e-mail: Karthika.d@sdnbvc.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_46

unreasonable quantity of defiled countries, for instance, the USA, Italy, Spain, and the overview go on. Regardless of the way that the traits of this novel coronavirus are not prominent, it can derive its lead from the experiences, for instance, the brisk pace of spread, weakness is higher for low safe, diabetic, and developed people, for the most part, phenomenal case of recovery rate and other information it has is that it spreads by a human-to-human transmission through globules or by direct contact, and the incubating period for the defilement has been estimated as 2–14 days. For instance, robotized and telerobotic structures significantly limit the risk of overpowering issue transmission to front line human administrations people through creation; it is possible to triage, evaluate, screen, and treat casualties from a protected division; a great deal occurrence of the detect is that the clinical, planning, and science systems get together to resource the restorative administration's structure, therapeutic administrations workers, and society; all through the propelled catastrophe are presented [2].

In development, amazingly, there is not either pharmaceutical for the settle or inoculation for neutralization starting on April 10, 2020 [3]. Nonetheless, nations have been trying treatment procedures and practice of various reasonable drugs (without legitimate confirmation). The preventive measures defined by the WHO are to rinse the hands each by and by and again with a chemical and water or with an alcohol-based hand rub, taking after social clearing, and rehearsing respiratory neatness. The cover will additionally expect work in shielding people from disease, and it is needed to recognize when and how to use the scatter. Beside it, the immaculate way the absolute best approach to stay aside from the ailment from spreading is by staying at home and keeping missing from the party, taking after social disengaging, and confining us from tarnishing individuals or individuals revealed two separates.

Headways have been constantly making bit by bit anyway during a bewildering condition as of COVID-19 across the dining table, crafted by these efforts to help humanity entirely unexpected procedures is great. The loyalty of these advances in fighting against novel crown illness might be either affecting genuinely impacting [4]. The past suggests to the progression by the performance of which the feast of the debasement is coordinated or the difficulties of the humanitarian organization's framework have been blasted well known. The last incorporates those headways that have been valuable in holding up the lawmaking body and everybody to distribute with the circumstance. The prime aim of our judgment is to investigate the distinctive development-based strategies that sponsorship the earth in controlling pandemic and besides other reasonable progressions that can be utilized. The effect of our assessment is to make the clients, specialists, and analysts to understand the duties of the developments in controlling the plague and further, to enable the use of the advances in not all that far away future.

2 Literature Review

Mohanta et al. [1], the perception things are linked with remote or wired for correspondence, handling, processing, and checking distinctive ongoing situations. The things are heterogeneous and have low retention, less handling power. The utilization of the IoT framework accompanies security and protection challenges because customary-based existing security conventions do not appropriate for its widgets. In this work, the creators at first depicted a diagram of the IoT innovation and the territory of its application. At that point, the creators efficiently study the three essential innovation machine learning (ML), artificial knowledge (AI), and blockchain for tending to the security issue with it. At long last, an interrogation of this overview, security issues unraveled by the ML, AI, and blockchain with research difficulties are noticing.

Chandra et al. [2], artificial insight is customized on PCs to portray human knowledge. It has gained tremendous publicity and has built up to reform pretty much every calling including lawful division. New legal recreated AI programming like Ross insight and Catalyst alongside machine learning and natural language processing gives feasible battle objectives, better authentic clearness, and better authorization to equity and new challenges to customary law offices offering lawful help using utilized partner associate model.

Makridakis [5] presumes that noteworthy upper hands will keep on accruing to those applying the Internet broadly and ready to face pioneering challenges to turn imaginative items/administrations into overall business examples of overwhelming hardship. The best test confronting social orders and firms would employ the advantages of benefiting AI advances, giving huge chances to both new items/administrations and gigantic efficiency upgrades while keeping a strategic distance from the threats and detriments as far as expanded joblessness and more noteworthy riches variations.

Yu et al. [4], artificial knowledge (AI) is a footstep by step changing clinical practice. With the ongoing advancement in digitized information procurement, AI, and processing framework, AI applications are venturing into territories that were recently thought to be just the territory of human specialists. In this review article, we plot late discoveries in AI innovations and their biomedical applications, recognize the difficulties for additional progress in clinical AI frameworks, and sum up the financial, legitimacy, and social ramifications of AI in medical services.

Peek et al. [6], a significant move from information based on information-driven strategies while the enthusiasm for another examination subject, for example, vulnerability the executives, picture and sign handling, and characteristic language preparing has been steady since the mid-1990s. These papers identifying with rules and conventions are among the most profoundly referred to. In concert, these two disciplines have given about 51% of everything being equal. There have been eight of these papers that were referred to in any event ten times each year since their dispersion.

3 Emerging Technologies for Handling Pandemic Situations

Creative strategies that lower back the general public are the one which has any kind of effect the people to take after confine, social separating reasonably in sort out to limit the spreading of the affliction and thus has any kind of effect in controlling the broad [6].

Figure 1 depicts situations such as healthcare, government, community, commerce, and vitality utilization which play a vital work in all circumstances in most cases as widespread challenges.

3.1 Quick-Thinking and Effective Pandemic on the Alert

Enormous estimations can permit checking of the contamination erupt constantly. With secure to past epidemics and pandemics scenes, COVID-19 is outstanding in that open-get section to datasets containing everyday amounts of new ailments hurt someplace close to the technique for the USA, and, in specific cases, even urban zones, are ordinarily available. Mixed in with the experiences, it causes around the action of individuals; it directs the proper real factors set to mix logical showing and AI [7]. Blue stain, a Toronto-based startup that uses an AI-unparalleled perception contraption, appears to have been the first to distinguish the scourge erupt, a few hours after its revolt inside the first suggested point of convergence of Wuhan, exact beforehand of the Chinese pros and different overall organizations and associations.

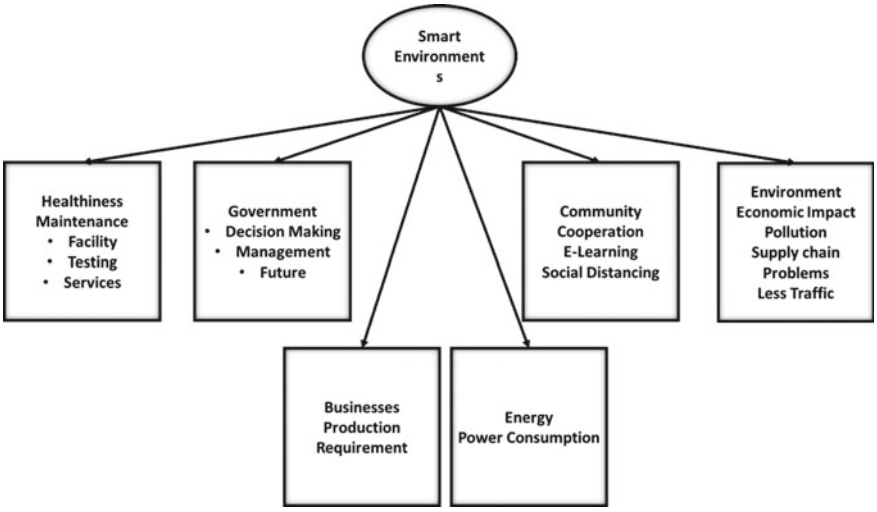


Fig. 1 Smart environments

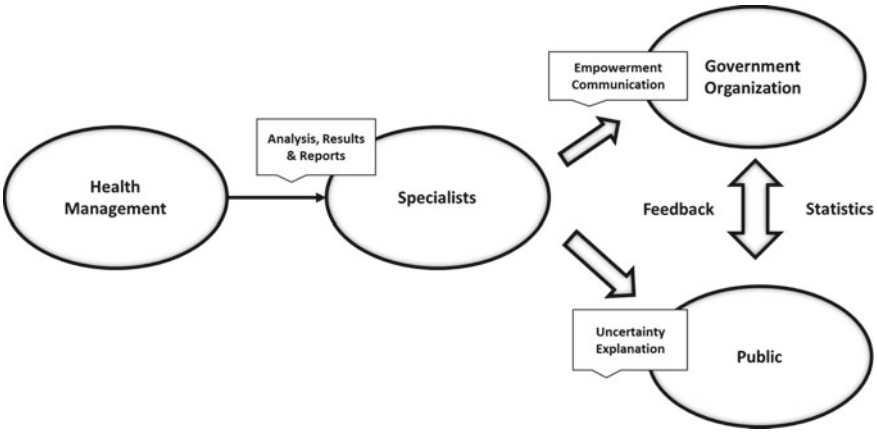


Fig. 2 Pandemic challenges in healthcare

3.2 Rapid Detection

A quick conclusion could be an indispensable part of human governments since it gives experiences concerning the therapeutic issue of a persistent and enlightens almost progressive social policy choices. In continuous decades, current strategies have been produced for the rapid diagnostics of different maladies [8]. The rule central purposes of quick disclosure methods are the peril of prior intercession and snappier associated with exercise to potential subjects, still in like manner extended throughput of the appraisal. Here, Fig. 2 will discuss confirming affirmation by using QRT-PCR, foundation screening by methods for smart adjusting administrator test, finally, AI-based picture getting ready methodologies using chest x-shaft/CT/HRCT pictures.

3.3 QRT-PCR

The most rising procedure for the abstract and quantitative ID of nucleic destructive from SARS-CoV-2 is higher and decrease. The breath models are a quantitative ongoing polymer chain response (QRT-PCR) and are in various events quantifiable all through the pointless zone of defilement [8]. QRT-PCR can trip as low as one contamination atom in swabs taken from the lip or nose. Test amassed from layout parts involving the patient's nose or throat is overseen different engineered courses of action that dispose of proteins and fatty tissues, leaving most first-rate concentrates of RNA within the mannequin.

3.4 Immunodiagnostic Test (Rapid Antibody Test)

The resistant symptomatic evaluations play a significant work for the convenient area of COVID-19 as saying earlier. This antigen area approach distinguishes the closeness of viral proteins (antigens) communicated by infers of the COVID-19 disease in a plant collected from the respiratory tract of a person [8].

3.5 Working from Home

In summation, a tainted character does not perceive whether the soul is contaminated because of the direction that the masses may be unwell with the infection running from 1 to 14 days before getting to the manifestations. On these lines, if the general public highlights not surprisingly, the opening will be at its top, and consequently, every association which would not genuinely need the nearness of representatives at the working patch is closed all through isolating period [9]. The basic requirements for a human or a lady to telecommute are a processing gadget, net openness, and cam if necessary. Meeting and gatherings can be executed remotely by way of a video meeting.

3.6 Wearables

Wearable development implies clever body-worn propelled contraptions for evaluating, examining, and communicating data. The real factors can be, for example, body pointers identified with all important signs and reactions and generous action. Wearable headways can moreover equip real factors to the wearer through exhibit science or contractible analysis. Standard models of customer wearable join interruption and vitals trackers that move people to use and effortlessly them with persistent comments on practice reasonability. The increased reality, and to a lesser extent AR advancement, is mechanically useful off-the-rack purchaser grade things. Social security workers can use progressed and AR advances to enable intuitive participation for a remote estimate or fix orchestrating [10].

3.7 Distance Learning

The COVID-19 that spreads rapidly has furthermore prompted us to turn off every single instructional organization. On these lines, understudies remaining at residential want to dab wearing some answer for continuing their learning. E-learning speaks to a progressive type of tutoring which gives rich acing experience [11]. E-learning

is electronic, mechanical expertise the spot sources can be given the guide of the assortment of means. The countries fundamental the graph of online tutoring incorporate the USA, UK, India, China, South Korea. Numerous stages have been created for separation becoming more acquainted with, for instance, Google Classroom, Coursers, Audacity, and so onward. In specific cases, the online meeting is remained gushed through video conferencing by the method of the teacher for a team of understudies. Every single of these advancements has randomly contributed in assisting individuals to keep on being at home.

3.8 Technologies Applicable for Pandemic Control

The former section offers upgrades that were commonly acknowledged to resist toward the novel crown contamination malady and were set apart as mechanical systems that were used to check the width and toughen the overall population [12, 13]. In this segment, sensible approaches and making effects that have the reason to help with encountering this far showing up at circumstance are talked about. These remarkable advances consolidate artificial experiences and machine learning.

3.9 Drone Technology

Automobiles are vehicles that affect utilizing a streamlined power that can be remotely manipulated. Those are self-ruling utilitarian roams which are being used within the military and the application [14]. This is coming out to industry, cultivating, improvement, open watching and perception, organic framework checking for any degradation, movement of items, and in clinical purposes.

3.10 Robotics Technologies

Mechanical innovation can aid permanently, and it may be shown improvement more than two or three respects. Nursing and restorative organizations enduring are work-centered work and mechanical progression internal that the human commitment's structure is an engaging form of motion [15]. With the toughening of artificial experiences and AI estimations, a mechanic can take from tremendous bits of knowledge and get full-size bits of information. These robots can be routinely positive in giving stop and fix when appears at a human parental figure. In advance, the adaptable robots have full-size central focuses in restorative organizations workplaces.

3.11 Mobile Applications

The improvement of information technology (IT) has made a huge platform in a genuine sum convenience and is speedily making inside seeing the web. Farther, the phones have cleared a way to give out with getting admission to these used sciences in an incredible arrangement significantly less extraordinary way. In the current world, mobile phones play a remarkable limit as they are embedded with perception phenomenal exhibits, outlandish choice cameras, enormous accessibility to the web, and support for a system of programming [16]. M-prosperity limits can nimbly legitimately help at whatever level and any area, thus, overcoming the land confinements, and it causes the therapeutic administrations people outfit perfect provider by methods for the utilization of mobile development.

4 The Outbreak: Challenges and Consequences

Terminate up to exact and solid records about the disorder, misdiagnosis of the novel crown infection as flu. The continuation of general undertaking and travel exacerbated the overall scatter of the taint. In the present, a segment of the cardinal elements that continue to perplex the limitation and alleviation attempts to incorporate a stretch of 2–14 days. In between the intro and start of appearances, the spread out of tainting by strategies for asymptomatic bearers [17]. In this bring forth period, constrained restorative administrations likely worldwide in the articulations of the advantageous extent of clinical association beds, showed staff, and clinical equipment to aid the characteristic casualties exhaustive of the fundamentally unwell events in concentrated thought units (ICUs). The troubling influence in the overall deftly chains incited by using the pandemic. Others, in the cost of ensuring the cognizance of fundamental organizations. For instance, organization, police and firefighting, water and power, food collecting and allocation, and transportation face a heightened danger of pollution as well.

5 Inference from the Review Work

Mechanized thinking gadgets are being investigated and used for key screening of feasible early pollutions of SARS-CoV-2 among individuals. Regarding the AI put together, getting data regarding structure, individuals are being arranged as high, moderate, or insignificant dangerous individuals. In whatever event, more than one authority differentiates as they might desire to thoroughly look at the data pulled together by methods for these AI instruments, as it is sometimes observed unstable owing to issues of across the nation secure. The chest CT-based finding, commonly

in COVID-19, has certain obstacles in articulations of partition between COVID-19 and system acquired pneumonia. In this manner, a significant getting data on the mannequin for COVID-19 used to be created and used to perceive the ordinary system picked up pneumonia from COVID-19. With a non-appearance of direction of care diagnostics, AI-driven devices can display important in finding the open door concerned and the transmission components among one-of-a-kind people social events. Ahead long, owing to the commitment of fiery learning strategies in AI instruments, the monetary value of trust in powerful methods has continued. Consequently, in that respect is a crushing need to standardize shows to make AI-based units that can be used all through such tragedy.

6 Conclusion

The existence of COVID-19 demands commonly affected all people over the earth and entirely modified the common faces of humanness. The resourceful methods that lead to reaching out to the disease compete across the world through communication. As a result, this article asserts that the knowledge of pandemic position primarily focuses on artificial intelligence and machine learning. These establishments involve AI, ML, and IoT which are to ramble on modernization and automation revolution. The growth of COVID-19 gets driven by the constraints of the improvements, and several alternatives have been recommended for understanding the infection. This got accelerated the evolutions which are happening in developments in quite a lot everyone's field. The practical knowledge that the people go around throughout this time might take a lengthy drag influence in an Internet shop, E-learning, automated payments, telehealth workplaces, and flexible entertainment. The upcoming effort will focus on additional innovative developing nanotypes of machinery to overwhelm problems that remain confronting in contagion circumstances.

References

1. Mohanta BK, Jena D, Satapathy U, Patnaik S (2020) Survey on IoT security: challenges and solution using machine learning, artificial intelligence and blockchain technology. Internet Things 100227
2. Chandra G, Gupta R, Agarwal N (2020) Role of artificial intelligence in transforming the justice delivery system in COVID-19 pandemic. Int J Emerg Technol 11:344–350
3. Elavarasan RM, Pugazhendhi R (2020) Restructured society and environment: a review on potential technological strategies to control the COVID-19 pandemic. Sci Total Environ 138858
4. Yu KH, Beam AL, Kohane IS (2018) Artificial intelligence in healthcare. Nat Biomed Eng 2(10):719–731
5. Makridakis S (2017) The forthcoming artificial intelligence (AI) revolution: its impact on society and firms. Futures 90:46–60
6. Peek N, Combi C, Marin R, Bellazzi R (2015) Thirty years of artificial intelligence in medicine (AIME) conferences: a review of research themes. Artif Intell Med 65(1):61–73

7. Wong ZS, Zhou J, Zhang Q (2019) Artificial intelligence for infectious disease big data analytics. Infect Dis Health 24(1):44–48
8. Kephart JO (2005) Research challenges of autonomic computing. In: Proceedings of the 27th international conference on software engineering, May 2005, pp 15–22
9. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR (2018) Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings. BMJ Glob Health 3(4):e000798
10. Asad F, Andersson M. Artificial intelligence for ecological sustainability—new machine learning applications for sorting of plastic waste?
11. Mehr H, Ash H, Fellow D (2017) Artificial intelligence for citizen services and government. Ash Center for Democratic Governance and Innovation, Harvard Kennedy School, Aug 1–12
12. Chin-Yee B, Upshur R (2019) Three problems with big data and artificial intelligence in medicine. Perspect Biol Med 62(2):237–256
13. Bendale S, Prasad J (2020) Preliminary study of software defined network on COVID-19 pandemic use cases. Available at SSRN 3612815
14. Hu Y, Jacob J, Parker GJ, Hawkes DJ, Hurst JR, Stoyanov D (2020) The challenges of deploying artificial intelligence models in a rapidly evolving pandemic. Nat Mach Intell 1–3
15. Allam Z, Dey G, Jones DS (2020) Artificial intelligence (AI) provided early detection of the coronavirus (COVID-19) in China and will influence future urban health policy internationally. AI 1(2):156–165
16. Radanliev P, De Roure D, Van Kleek M (2020) Digitalization of COVID-19 pandemic management and cyber risk from connected systems. arXiv preprint arXiv:2005.12409
17. Crump C, Schlachta-Fairchild LM (2020) Achieving a trusted, reliable, AI-ready infrastructure for military medicine and civilian trauma care. In: Artificial intelligence and machine learning for multi-domain operations applications II, vol 11413, Apr 2020. International Society for Optics and Photonics, p 114130C

Interpretation of Handwritten Documents Using ML Algorithms

Mohd Safiya, Pille Kamakshi, Thota Mahesh Kumar, and T. Senthil Murugan

Abstract Handwritten character recognition is a continuing field of research which covers artificial intelligence, computer vision, neural networks, and pattern recognition. An algorithm that executes handwriting recognition can acquire and detect characteristics from given handwritten document/image as input and convert them to a machine-readable form. It is also known as the task to convert the input text to extracting the features with the help of symbol representation and icons of each letter. The main goal of this handwritten text recognition is to identify an input character and text on a scanned image which are written in cursive writing, and these features are extracted with each input character pixels. Each character dataset contains 26 alphabets. IAM datasets are used for training the characters and for classification and recognition. The output is generated in the form of human-readable text. In this paper, the handwritten documents has been interpreted using machine learning algorithms such as connectionist temporal classification (CTC), long short-term memory networks (LSTMs), and generative adversarial networks (GANs). The results show the comparison of feature extraction based on proposed approach with convolutional neural networks (CNN) and recurrent neural network (RNN) with real-time datasets gain in better performance when using these algorithms.

Keywords Classification · Feature extraction · Segmentation · LSTM · GAN

1 Introduction

The strength of a computer device or a software program is to analyze the shape of cursive writing on documents or on papers that are translated in the form of letters, words, alphabets, sentences, and punctuation marks, etc. Handwritten text recognition is the conversion of text which is written into the digital form where it can lift up the pen tip movement as well as moving the pen up and down switching modes for a cursive text. Handwriting recognition systems handle formatting the text,

M. Safiya (✉) · P. Kamakshi · T. M. Kumar · T. Senthil Murugan
Kakatiya Institute of Technology and Science, Warangal 506015, India
e-mail: safiamohd241@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_47

segment the character, and find the correct words. The goal of this handwritten text recognition is to identify the input characters on the image which are written in cursive writing, and these features are extracted with each input character pixels. Handwriting recognition is used for digital libraries allowing for image textual information into different stages or modules such as preprocessing, segmentation feature extraction, and classification by using the machine learning algorithms.

In this present situation, many researchers had worked with several types of machine learning algorithms like support vector machine (SVM), random forest (RF), k-nearest neighbor (KNN), decision tree, convolutional neural network, other neural networks, and many other algorithms. Researchers combined different types of machine learning approaches and algorithms with an image processing method to increase the classifier with a best accuracy measure for the handwritten characters and text recognition system. In this paper, the hand written documents has been interpreted using machine learning algorithms gives good performance evaluation given by the deep learning structure considered like connectionist temporal classification (CTC), long short-term memory networks (LSTMs), and generative adversarial networks (GANs). The results show the comparison of feature extraction based on proposed approach with convolutional neural networks (CNNs) and recurrent neural network (RNN) with real-time datasets gain in better performance when using these algorithms.

1.1 Deep Learning

Deep learning is nothing but the training model with machine learning and artificial intelligence which can think like a human brain and can implement input data image to predict the output by using high performance in data abstraction, and planning structure is composed of more than one nonlinear transformation. Deep learning requires very large datasets, and it is computationally intensive.

It learns features and classifiers automatically, and it gives unlimited accuracy. Deep learning is used in this approach because the handwritten text should be in different cursive writing with an IAM dataset to extract the features from an image, classify automatically, and generate the output in the form of normal human-readable text with the help of proposed algorithms for gaining better accuracy.

This approach includes the different algorithms such as follows.

1.2 Long Short-Term Memory (LSTM)

Long short-term memory (LSTM) is used for sequence of linear-order text in an artificial neural network and also in the concept of deep learning architecture. It is not similar as to standard feed forward neural networks. LSTM has feedback connections for the length of the data. It cannot process the one data point at a time,

and it can process the sequences of continuous data. For example, LSTM is used for a task like handwriting text, character recognition, speech recognition, and voice recognition.

1.3 Generative Adversarial Networks (GANs)

Generative adversarial network (GAN) is an unsupervised learning in machine learning algorithms that automatically involves different learning and discovers the regular expressions and patterns in input data. Recently, generative adversarial networks show good performance on image generation and style transfer of images. However, GANs do not work better with text on real-valued continuous data generation, where texts should be only discrete data. GANs are the model architecture for training a generative model. For example, GANs are used in video games, animated movies, or cartoons.

1.4 Connectionist Temporal Classification (CTC)

In this classification, connectionist temporal classification (CTC) algorithm is used to give training for the neural network models. It may similar as loss function calculation using cross-entropy. It is used in some problems where having aligned data is an issue like speech recognition and text recognitions. The CTC algorithm is an alignment or adjustment free algorithm, and it does not require an indentation between the input and the output text. It can be used for tasks to perform like online handwriting text recognition and speech recognition audio. CTC refers to the results and scoring and is independent of the basic neural network structure.

2 Related Work

Yang et al. [1] proposed text independency writer identification which is increased in real-time applications for general and efficient way, and the proposed framework is used in the RNN approach which deals with the raw information directly and online handwriting and extraction of the features with RHS representation.

Rehman and Khan [2] introduced the recognition of Urdu online character recognition which is an alternative problem to perform scaling, rotating, orientation, different text styles of writing, and the small number of text that are used for model training for the datasets. In this approach, many methods and performance measures like specificity, accuracy, confusion matrix, Mathews correlation coefficient (MCC), Receiving operating characteristics (ROC) and precision are computed and evaluated the proposed model.

Liu et al. [3] proposed an outlined little bit of recent works that is related to DCNN-based methods and techniques on OLHCCR and OLCHH. The SSDCNN combines a static feature through a deep neural network and fully connected neural network and also uses the eight directed path features, where the feature paths have been proved by an efficient way in OLHCCR and SSDCNN preserves the strength of the probability features.

Farooqui et al. [4] introduced an approach for performing an assignment for document retrieval like word spotting-based methods available through the recent online character recognition, where the system is to recognize the text of each word and data generation methods.

Memon et al. [5] proposed the methods like a structured model that can be divided into small groups on the specific information structure such as like graphical approach and grammar-based approach. The image should be in the format of binary with RGB structure, and boundary is considered in a perfect condition applied for the OCR.

Yu and Saniie [6] introduced a technique for generating GAN architecture that had implementation in morphology with an image and image style transferring the image styles from one to another style. It had a good content in this field. The success should be very useful and helpful to handwritten text which is the approach of generating the text in GAN algorithm.

3 Implementation

In this implementation section, the interpretation of handwritten documents using machine learning algorithms includes CTC, LSTM, and GANs, where the algorithms are used to perform with best accuracy and measures of predicated values while training and testing the datasets. This approach includes different methods for recognizing the cursive text into normal text implementing with the methods that are as follows.

3.1 Preprocessing

Preprocessing the data for classification is to clean the information in order to reduce the noise, handle the missing values, and also remove the irrelevant and redundancy attributes and information. In preprocessing stage, it improves the performance and enhances the quality of the handwritten scanned text image for next methods for implementation. It can be highlighted and can be divided into heavy small tasks, like paragraph text detection, text segmentation, word segmentation, character segmentation, image normalization, and many more.

3.2 Segmentation

In this step, the character set will be identified as letters, and the text image will be converted to normal text. This approach is based on system mainly used especially when the text is separated between lines and characters that are difficult to be determined, like difficult background, text lines, and every character in the image. Segmentation is to divide and group specific pixels from a character and numeric image, assign them labels, and classify further pixels according to the labels.

3.3 Feature Extraction

Feature extraction is a variety method of reducing dimensionality and pattern recognition in an image processing. When the input information for algorithm is very heavy to be processed which is an unnecessary thing, the input text will be transferred into less representation set of features, and it should be called as a feature vector. Transferring the scanned text into the set of the features is called feature extraction. In the feature extraction method, we have to extract the featured image pixels from scanned image to gain the features that can be used for measuring the high classifier performance measure of each classification model.

3.4 Classification

In the classification stage, here we are doing the train and test the deep neural network with different algorithms. In this stage, the features are the input, which is a trained classifier, where the training model classifier classifies the prediction of each and every characters and words. Some methods are applied to recognize the scanned image with data and generate the text, in which the image is subdivided into segmentation approach, and segmentation approach is used for recognizing image with the text.

4 Results and Discussion

In this approach, the deep learning algorithms are used with abovementioned four steps for handwriting text recognition. Firstly, we have to import the packages, read the A_Z handwritten comma separated values (CSV) file dataset, split the data read into the images and their corresponding labels, then reshape the data in the CSV file so that it can be displayed as an image, after that plot the numbers of alphabets in the dataset, and after that we shuffle some of the images data of the train set. Data reshapes

the training test dataset so that it can be put in the model. After putting into the model, the deep learning algorithms are applied for compiling and fitting the model and getting the train and validation accuracies and loss with number of epochs which gives the good performance among those three deep learning algorithms. IAM datasets for English documents and images are extracted from the Lancaster-Oslo/Bergen (LOB) corpus, which had been taken by different researchers. The dataset that includes 1199 texted images with segmentation, in which 459,308 characters are distributed into text lines with 10,373 and symbols are 79 and so on. Following is the example text for the handwriting text recognition that should be converted into the normal human-readable text. A sentence and the words are chosen that are shown in the image in Fig. 1. The text for the sentence is as follows: "converting the cursive text to normal text using handwritten". In the sentence, there are nine words, 52 total characters, the space is excluded, and 58 total characters are included with the space in the given sentence as shown in Fig. 1.

Above is the cursive writing text that should be converted into normal human-readable text using the deep learning algorithms. This works on three deep learning algorithms, LSTM, CTC and GANs; the accuracy of each technique as for GAN is 82.6%, LSTM is 78.10%, and CTC is 73.55%. Few examples from IAM dataset which is converted into cursive image to normal text area are shown in Fig. 2.

Converted text: Handwritten text recognition using machine Learning techniques in applications of NLP.

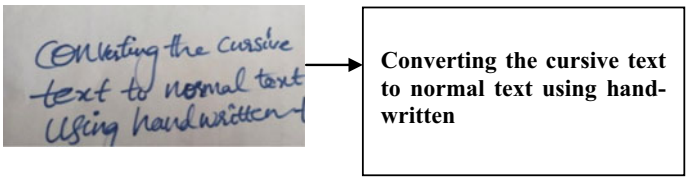


Fig. 1 Original and generated text

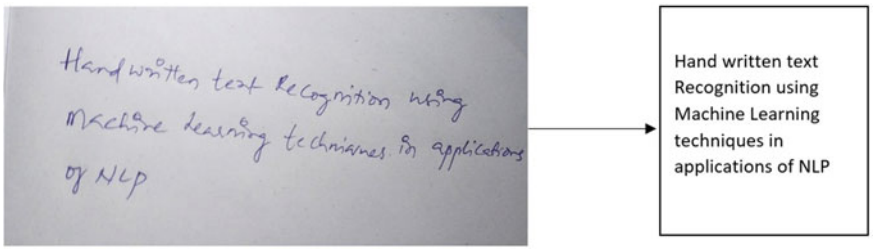


Fig. 2 Original and generated text

4.1 Comparison of Different Algorithms

In this paper, Deep learning techniques as GAN, LSTM, and CTC have been implemented and compared the results. Among these three algorithms, the GAN can perform well in normal text generation from handwritten images. This algorithm only produced 82.7% in accuracy and it is the highest one among three algorithms. The LSTM algorithm produced 78.10% in accuracy and CTC algorithm produced 73.55% in accuracy. All three algorithms are used same IAM dataset.

5 Conclusion

In this approach, the text recognition is used to learn and study about the English letters and symbols and many more. The handwritten letters are complex to learn, but from the present of characters, symbols and shapes with different diagrams of more letters are same. This approach is mainly helpful for the work toward other unknown scripts to convert normal text. Also as the architecture of deep learning neural network of GAN is for making, it will be able to encode text input with generate cursive handwriting images. The CTC algorithm is an alignment-free and indentation condition. It does not require an alignment between the output and the input of a given text. The LSTM cannot process the single data points but can also process the sequences of data. It should compare both CNN and RNN approaches and perform better accuracy.

References

1. Yang W, Jin L, Liu M (2016) DeepWriterID: an end-to-end online text-independent writer identification system. IEEE Intell Syst 31(2):45–53. https://doi.org/10.1109/MIS.2016.22
2. Rehman KUU, Khan YD (2019) A scale and rotation invariant Urdu Nastalique ligature recognition using cascade forward backpropagation neural network. IEEE Access 7:120648–120669. https://doi.org/10.1109/ACCESS.2019.2936363
3. Liu X, Hu B, Chen Q, Wu X, You J (2020) Stroke sequence-dependent deep convolutional neural network for online handwritten Chinese character recognition. IEEE Trans Neural Netw Learn Syst 31(11):4637–4648. https://doi.org/10.1109/TNNLS.2019.2956965
4. Farooqui FF, Hassan M, Younis MS, Siddhu MK (2020) Offline hand written Urdu word spotting using random data generation. IEEE Access 8:131119–131136. https://doi.org/10.1109/ACCESS.2020.3010166
5. Memon J, Sami M, Khan RA, Uddin M (2020) Handwritten optical character recognition (OCR): a comprehensive systematic literature review (SLR). IEEE Access 8:142642–142668. https://doi.org/10.1109/ACCESS.2020.3012542
6. Yu X, Saniie J (2020) Recognizing English cursive using generative adversarial networks. In: 2020 IEEE international conference on electro information technology (EIT), pp 293–296. https://doi.org/10.1109/EIT48999.2020.9208298

Survey of Streaming Clustering Algorithms in Machine Learning on Big Data Architecture

Madhuri Parekh and Madhu Shukla

Abstract Machine learning is becoming increasingly popular in a range of fields. Big data helps machine learning algorithms better timely and accurate recommendations than ever before. Machine learning big data phases and subsystems point the way to comparable issues and threats, as well as to associated research in various of previously new of unexplored areas. In latest days, data stream mining has become a popular research topic. The biggest challenge in streaming data mining is extracting important information directly from a vast, persistent, and changeable data stream in just one scan. Clustering is a powerful method for resolving this issue. Financial transactions, electronic communications, and other fields can benefit from data stream clustering. This research examines approaches for synthesizing enormous data streams, including projection clusters for high-dimensional data, scalability, and spreading computing, as well as the issues of big data and machine learning. The data abstraction and regular features of learning streams, such as abstraction wander, data flow system, and outlier observation, are discussed in this study. AutoCloud, which seems supported by previous inform prospect of typicality with eccentricity data science, will be used to discover deviations but instead resolve them using the support clustering approach, which is the process of computing difficulty and clustering exactness, will be detailed in this article. We present the MLAutoCloud algorithm, which will be used in machine learning PySpark frameworks. Implement the MLAutoCloud algorithm in the future to tackle the AutoCloud algorithm's problems.

Keywords Clustering · Big data analysis · Machine learning · Framework · PySpark · Stream computing

M. Parekh (✉)
Computer Engineering Department, Ph.d scholar Marwadi University, Rajkot, India
e-mail: madhuri.parekh109023@marwadiuniversity.ac.in

M. Shukla
Computer Engineering Department, Associate Professor & Head Computer Engineering-AI, Marwadi University, Rajkot, India
e-mail: madhu.shukla@marwadieducation.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_48

1 Introduction

In fields as diverse as computer visualization, audio giving out, natural language comprehension, neuroscience, strength, and the Internet of Things, machine learning (ML) approaches have had enormous societal consequences. The coming on of the large data period has sparked widespread interest in machine learning [1]. Big data has never promised or challenged machine learning techniques to obtain fresh insights into a variety of corporate applications and individual behavior [2]. Users access machine learning approach to identify internal structure and generate predictions from huge datasets. Machine learning thrives on strong computer environments, efficient learning approaches, with rich and/or huge data. Big data, for example, enables pattern discovery at various levels of granularity and diversity from multiple perspectives in an essentially simultaneous manner [3]. The focus of the research is on machine learning approaches in relation to huge data and current computer systems. We want to look into the benefits and drawbacks of machine learning on huge data. Big data opens up new possibilities for machine learning [4]. We offer a structure for machine learning on immense data in this research to better guide the discussion of its opportunities and problems.

2 Big Data and Machine Learning Framework

Figure 1 depicts the framework in favor of machine learning on large information (MLBiD). Machine learning is at the heart of MLBiD, and it interacts by means of four additional mechanisms: big information, client, domain, and scheme. Interactions take place in both directions [4].

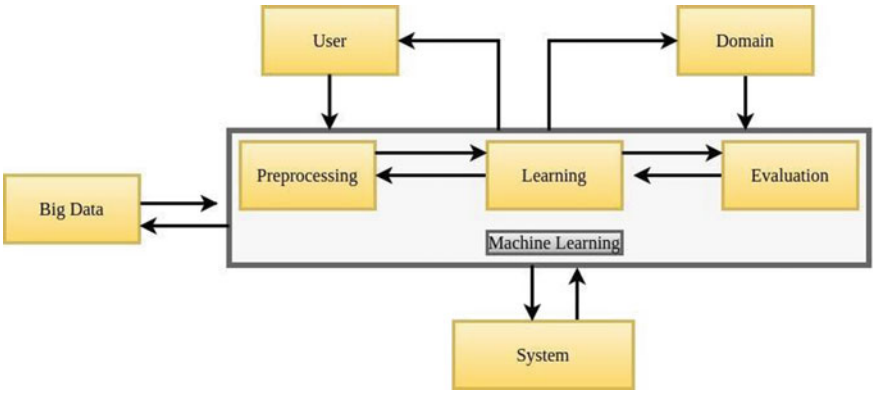


Fig. 1 Machine learning framework with big data (MLBiD) [4]

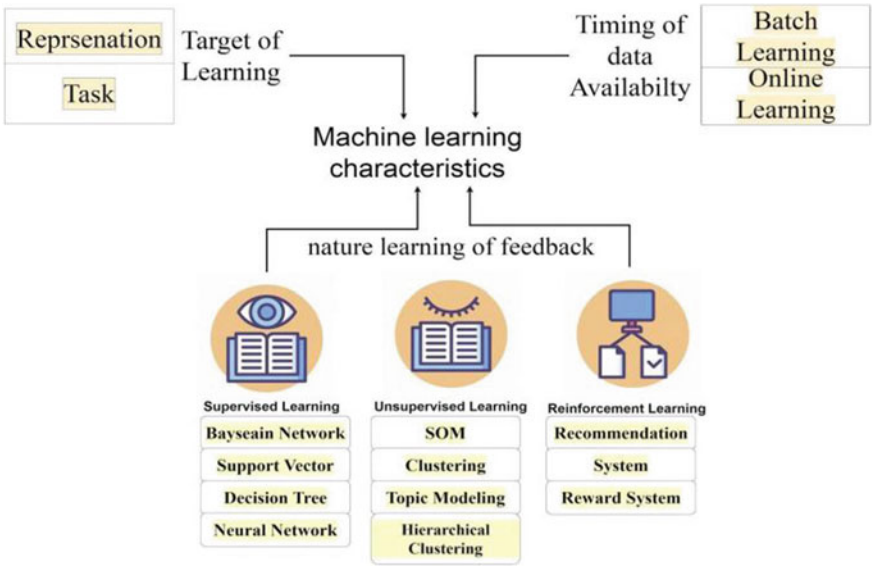


Fig. 2 Categorization of ML [4]

2.1 Machine Learning

Preprocessing stage aids in the transformation of unrefined data into the "correct shape" in favor of further knowledge procedures. It is likely that the unrefined data is shapeless, loud, partial, and unpredictable. Through data cleaning, extraction, transformation, and fusion, the pre-giving out step turns such information into an outline with the purpose of may be second-hand as inputs to learn [4]. As a result, as illustrated in Fig. 2, we consider a new categorization of ML.

2.2 Big Data

Volume its means extent or quantity of data, velocity its means rapidity of information invention, diversity its means category, environment, and arrangement of information, and veracity its means responsibility or excellence of acquire information are the four key scope of big data in Fig. 3.

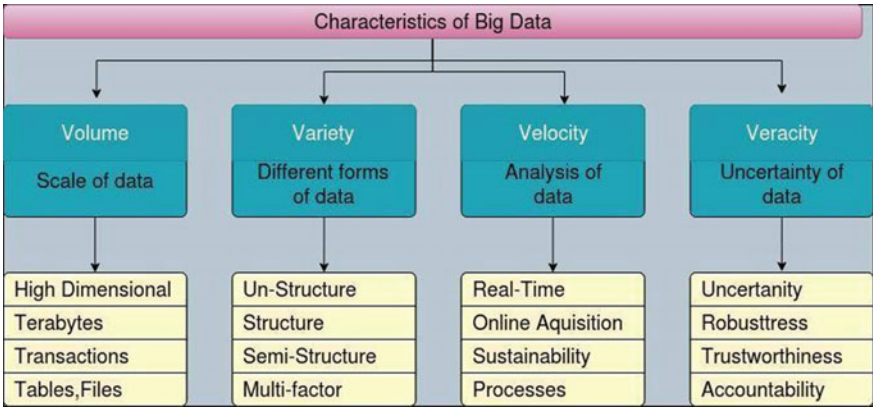


Fig. 3 Elements of big data [4]

2.3 Users, Domains, Systems

User interaction has been limited to provide that information label, answer domain-related question, or provide comment on learnt findings, which is often mediate by practitioners, resulting in extensive and asynchronous iterations [4]. Domain knowledge aids machine learning in detecting interesting patterns that might otherwise be missed by looking at datasets alone. The proposed system, often known as the platform, is a collection of programs that create and background inside which machine learning algorithms may execute. A multicore computer by means of disseminated design, for example, be predicted toward increases ML effectiveness as compared to simpler equivalents [4]. To solve the problems of big data, new tools and system architectures, such as hadoop or spark, contain been planned.

3 Challenges and Opportunities

Base on top of the MLBiD structure, we have included identified key opportunity and main challenges [5].

3.1 Possibilities and Challenges of Data Preparation

Data preparation tries to solve problems such data redundancy, discrepancy, sound, heterogeneity, revolution, labeling, information inequity, and quality representation or collection [5]. Conception flow is defined as change in the uncertain allocation of the desired production specified the contribution, in spite of the fact that the contribution distribution may remain static [6]. This becomes a difficulty when machine learning model is constructed using grown-up information and no longer appropriately respond to the sharing of fresh data. The representation is built by resources of only the sample starting the preparation gap, which is enthused to include those majority current examples, which is a possible approach to handle with concept drift. The windowing method implies that the most current data is the most important. Concept drift can take several forms: incremental, steady, rapid, and frequent all with its own set of difficulties. Data stream is frequently generated in non-stationary situations with rejection previous knowledge of the information allocation or real cluster count. Furthermore, the quantity of clusters and the distribution of data may alter over moment in time [7]. As a result, unlike batch techniques, algorithms for cluster online data streams procedure input single component at a time, update cluster structures as needed, and keep a little amount of matter or prototype in memory as a representation of the cluster. A gap exists in the stream partitioning-based technique, and instance complexity is small. It lacks the versatility of computer clusters at a subscriber span of time [8]. Clu-stream separation and structured gives the ability to generate clusters at a subscriber particular time. Keep the theme from drifting and give the cluster of random shapes [9]. High-dimensional, predictable streaming data cluster using HPStream splitting and complex-based clustering. A standard measurement for a common digit [10]. DBSCAN's thickness-based capability of scrutinizing the cluster for massive data makes parameter setup extremely tough [11]. Cluster formation is plotted using a stream-optics density-based clustering approach based on time. It is not a supervising approach for cluster pulling out [12]. The MR-Stream density-based clustering method increases clustering procedure. It is lacking in elevated spatial information. It appears to be functioning [13]. The HDD-Stream density-based clustering approach clusters information with a higher

spatial resolution. In order to find adjacent clusters, it takes an additional amount of time [14]. The supply of latest often these values is precisely captured employing the DENGRIS density-based clustering method with down windowpane representation [15]. The SOMKE clustering algorithm makes efficient and imaginative use of non-stationary data. It cannot hold information that has been distorted [16]. For actual data streams, the LeaDen-Stream density and grid-based clustering algorithm assists thickness decomposition and detects rising activity. It does not keep track of different dimensional data [17]. The POD–Clus model-based clustering method helps with data disappearing and conception evolution. In data streams, the computation and updating of pair wise space takes a lot of occasion [18]. If the cluster is not in a spherical shape, the BIRCH hierarchical-based clustering method overcomes the inability to undo what was done in a previous step [19]. The clustering algorithm based on SPE-cluster partitioning cannot find arbitrary outline clusters [20]. The outliers are efficiently identified using a Cobweb model-based clustering algorithm. It does not provide a compressed representation of the data [21]. Clustering is a crucial component of data mining. Because it can identify arbitrary form groups and outliers, the density-based clustering [22] method is appropriate for grouping data streams [23]. A cluster is defined as area of higher thickness than the rest of the information set in density-based clustering. The detection of outliers in clustering methods necessitates highly time-consuming calculations. Conduct noisy data, restricted moment in time and memory, changing data, and high-dimensional information must all be taken into account [24]. An outlier [25] is a in-order point that exhibits aberrant behavior in the system and is reliant on the application [26]. Clustering is a technique that may be used to solve problems like these in stream data mining.

3.2 Related Work

Zhou et al.—Machine knowledge on big Data openings and Challenges [27]. In this paper, ML for important data concentrated on the volume, velocity, and variety aspects. To cope up with data integrity, one suitable approach is to design tools that can assess the responsibility or authenticity of information or datasets, allowing untrustworthy data to be screened out during data from before the; the other is to establish protection Ml algorithms that can infer with untrusted or even conflicting data. Massive online information for data streams [28] will be discussed henceforth. With inaccurate or non-concurring information, exacting data veracity knowledge is required. Extracting and compression large amounts of data. Selection of scattered points on a large scale. Application-aware and process management. For just a wide range of multimedia material, multi-view expertise is essential. Multimedia neural semantic embedding is a type of multimedia neural semantic embedding. Machine Learning With Big Data Challenges and Approaches, by L'Heureux et al. [26]. Clustering is the best method for analysis of stream data. As labeling of data is not possible for streams so clustering may assist this process [29]. In this research, there will be a systematic evaluation of the problems related to machine knowledge

in the distributed environment, as well as a categorical evaluation of the volume, diversity, and verity extent of big data. Inquire about optimistic future advances in machine learning using Large Data sets. for further discussion the with assist of computer science and Big Data Kosta et al. A modified qualified OPTICS Algorithm Approach for Data Methodologies Research in Engineering, Technology, and Applied Science Research. The channel optics algorithm is related to a flexibility of pruning in this research, as well as the detection of true outliers, which reduces memory consumption and boosts the speed of linking implicit clusters [30]. In an early investigation, generality drift is used to generate clusters of arbitrary shapes [31].

4 Existing Model Problem Survey

The method of pulling out information stream by proposing data cluster object in a particular given situation is due to various fundamental principle: (i) single examination cluster put data clustering has get to be done rapidly conscionable one incident, successful a particular go by due to the information stream incoming constantly; (ii) incomplete instance put data cluster have to be shaped in real time within an incomplete instance framing; (iii) with just incomplete memory, the clustering technique is made available, but it is to method acting constant, conventional, unrestricted material stream; (iv) indefinite quantity and conformation of clusters: These features of the information flow stay unexplored before process; (v) developing information put the algorithmic rule has got to be considered in so much how on be organized to hold the ever-altering aspect of the information flow. To take away some of the beyond challenge, use the AutoCloud algorithm in data flow [29, 32]. It is used for difference detection mission which is based on the presently introduced common idea of typicality and eccentricity data analytics. It permits development and coming collectively of cluster self-resolute as new data combined becomes accessible. AutoCloud does not require preparation or previous information about information set, and it has structure without predefined shapes or limitations. These techniques are developing, online, and recursive method. AutoCloud uses fuzzy logic concept which permits each data section to be connected with a number of data clouds concurrently. AutoCloud is based upon the concepts of data streams. Concept flow and concept of development are in addition handled by AutoCloud algorithm. These concepts are inbuilt in data stream algorithms in general which open do research areas in AutoCloud algorithm [32]. It will raise the necessity for exploration of work on the algorithm which would cater the difficulty and resolution of perception of flow and evolution.

4.1 Compensation of AutoCloud Algorithm [32]

It compensates ever-evolving clusters as tiny structures and information cloud, where non-single parameter of all grain can be taking, but also new cluster can be imaginary and present same integrated, which kind it appropriate for managing dynamical and developing information which is perception flow and development. It does not involve accumulate preceding information sample in remembering, execute accumulation dispensation work, or using sliding windows, so it is iterative calculation. Applied science is completely unsupervised because it does not need offline formulation or preceding knowledge about the info and can be in advancement from an unfilled knowledge support.

4.2 Drawback of AutoCloud

Because information clouds have no specific shape or bound and are not directly derived from the type of compatible analysis implemented, these are visually represented as ellipsoids in understandable algorithms. An obtainable algorithm, that is, AutoCloud investigates use of the Euclidean distance which gives the clusters shape as ellipses which practically forces the meaning of cluster. In the AutoCloud algorithm, it cannot explore how to discover out the adaptive importance of m (Fig. 4).

5 Proposed Work

Incoming this manuscript, we specify an algorithmic rule for data stream cluster permitted MLAutoCloud. It will not require any explicit skillful model or any preceding knowledge for the information set to be synthesized. The projected in this manuscript is the MLAutoCloud algorithm which is completely data-driven. The perception of typicality and concentricity information analytics presents conception of typically helpful to deviation finds difficulty which is supported on Mahalanobis distance advice capable strategy to split clouds dynamically in the MLAutoCloud algorithm. This manuscript proposes to obtain an adaptive value of m which straight affects the compassion of the variance detector. MLAutoCloud algorithm will execute PySpark architecture which solves the challenge of big data flow rapidity.

5.1 Benefit of MLAutoCloud Algorithm

MLAutoCloud algorithm will be going to examine the use of the Mahalanobis reserve or cosine reserve element a metrical and to intend capable strategy to fragmented

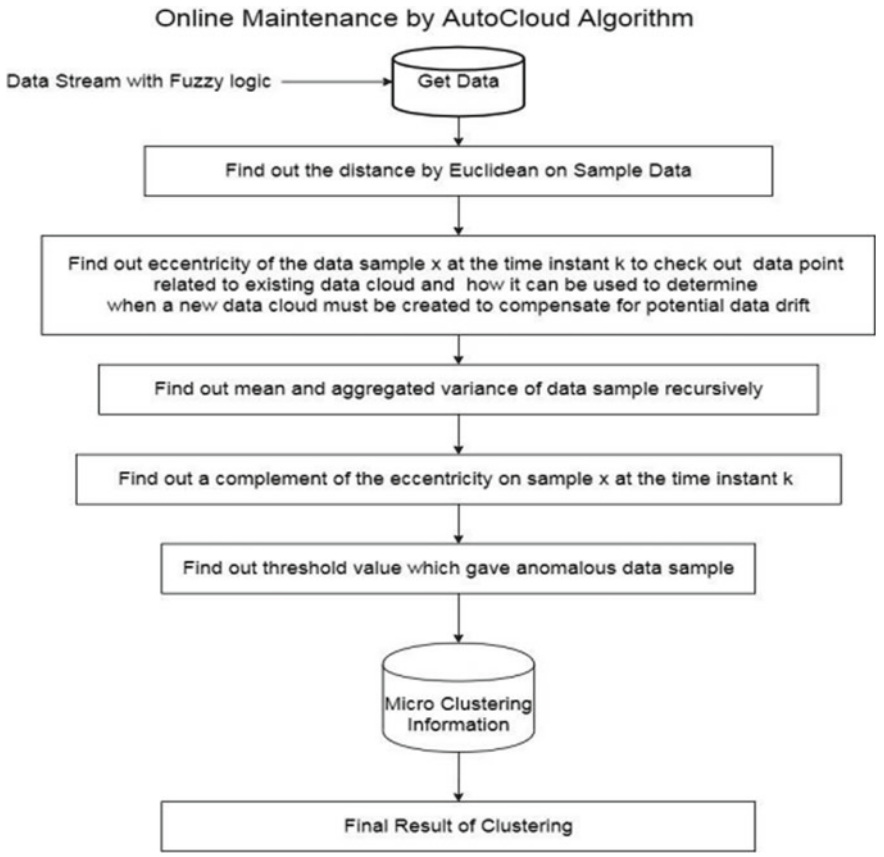


Fig. 4 Flow of AutoCloud algorithm [32]

clouds changing. In addition, MLAutoCloud aims at investigating how else to obtain an adaptable value of m, with this price becoming more efficient with each subsequent model analysis. The constraint m is user-defined and impacts the variation detector's compassion inside a single direction. MLAutoCloud online stream will apply PySpark architecture which solves the challenge of rapidity in big data stream (Fig. 5).

6 Conclusion

This paper provides a comprehensive evaluation of stream clustering algorithms with their challenges. In the next research work, a detailed study of big data with the above-said frameworks will be done. The investigation looks at how AutoCloud's available algorithms compensate for the difficulties of evaluating the difficulty of incomplete

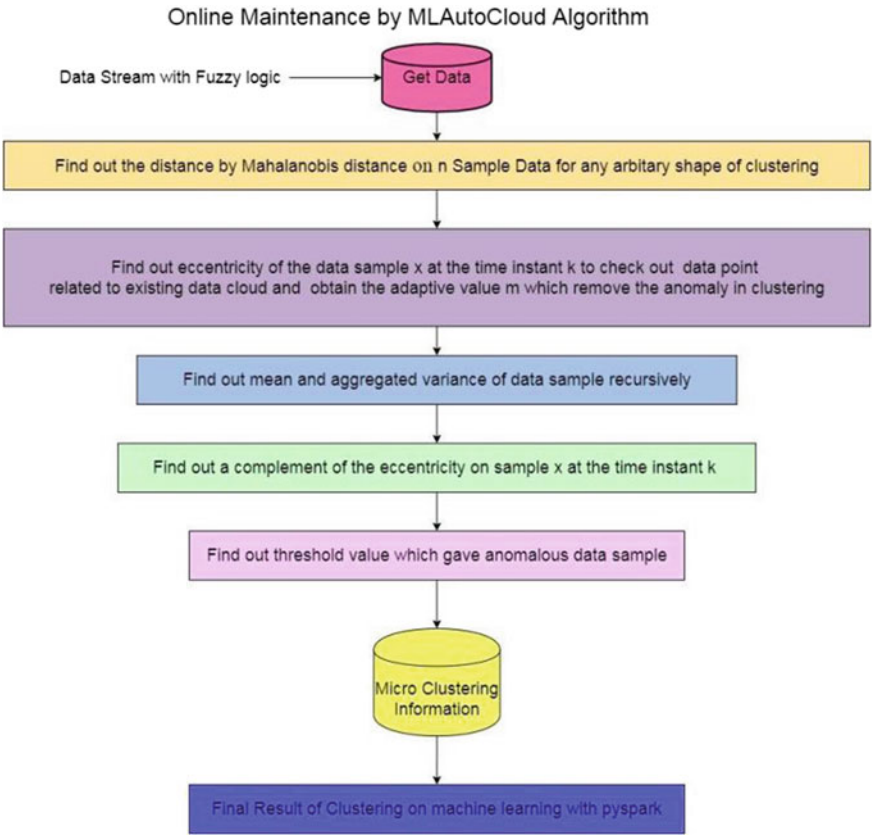


Fig. 5 Flow of MLAutoCloud algorithm [32]

memory using a particular technique. Anomalies in data flow are also hidden, but clustering algorithm can detect outlier and noise. AutoCloud algorithms, on the other hand, must keep conception drifts under check and detect uninform, shaped clusters. The phases of the AutoCloud method that have significant challenges will be solved by proposing the MLAutoCloud algorithm in this article. In the upcoming, we will offer the MLAutoCloud method, which is based on the TEDA machine learning idea integrated with big data architecture. Furthermore, in the future, the MLAutoCloud method will resolve to apply any random shape of clustering and will also acquire the adaptive value of *m*, which will eliminate differences, and will indicate the possible stream advance to the cluster through PySpark architecture.

References

1. Mavragani A, Ochoa G, Tsagarakis KP (2018) Assessing the methods, tools, and statistical procedures in google trends research: systematic review
2. Kolajo T, Daramola O, Adebiyi A (2019) Big data stream analysis: a systematic literature review. J Big Data
3. Qian Z, He Y, Su C, Wu Z, Zhu H, Zhang T, Zhou L, Yu Y, Zhang Z (2013) TimeStream: reliable stream computation in the cloud. In: Proceedings of the eighth ACM European conference on computer systems, EuroSys 2013. ACM Press, Prague, pp 1–4
4. Zhou L, Pan S, Wang J, Vasilakos AV (2017) Machine learning on big data: opportunities and challenges
5. Tsai C-W, Lai C-F, Chao H-C, Vasilakos AV (2015) Big data analytics: a survey. J Big Data 2:1–32
6. Russell S, Norvig P (2010) Artificial intelligence: a modern approach, 3rd edn. Prentice Hall
7. Bengio Y, Courville A, Vincent P (2013) Representation learning: a review and new perspectives. IEEE Trans Pattern Anal Mach Intell 35
8. O'Callaghan L, Mishra N, Meyerson A, Guha S, Motwani R (2002) Streaming-data algorithms for high-quality clustering. In: 18th international conference on data engineering, pp 685–694
9. Aggarwal CC, Han J, Wang J, Yu PS (2003) A framework for clustering evolving data streams. In: International conference on very large databases, vol 29, pp 81–92
10. Aggarwal CC, Han J, Wang J, Yu PS (2004) A framework for projected clustering of high dimensional data streams. In: Thirtieth international conference on very large data bases, vol 30, pp 852–863
11. Dharni C, Bnasal M (2013) An enhancement of DBSCAN algorithm to dissect cluster for large datasets. In: IEEE international conference on MOOC innovation and technology in education (MITE), pp 42–46
12. Ankerst M, Breunig MM, Kriegel H, Sander J (1999) OPTICS ordering points to identify the clustering structure. ACM SIGMOD Rec 28(2):49–60
13. Wan L, Ng WK, Dang XH, Yu PS, Zhang K (2009) Viscosity-grounded clustering of data aqueducts at multiple judgments. ACM Trans Knowl Discov Data (TKDD) 3(3):1–28
14. Ntoutsi I, Zimek A, Palpanas T, Kröger P, Kriegel H (2012) Viscosity-grounded projected clustering over high dimensional data aqueducts. In: Society of industrial and applied mathematics (SIAM) international conference on data mining, pp 987–998
15. Amini A, Wah TY (2012) DENGRIS-Stream: a viscosity-grid grounded clustering algorithm for evolving data aqueducts over sliding window. In: International conference on data mining computer engineering, pp 206–211
16. Cao Y, He H, Man H (2012) SOMKE kernel viscosity estimation over data aqueducts by sequences of tone-organizing charts. IEEE Trans Neural Netw Learn Syst 23(8):1254–1268
17. Amini A, Wah TY (2013) LeaDen-Stream: a leader density-based clustering algorithm over evolving data stream. J Comput Commun 1(5):26–31
18. Rodrigues PP, Gama J, Pedroso JP (2008) ODAC: hierarchical clustering of time series data streams. IEEE Trans Knowl Data Eng 20(5):615–627
19. Zhang T, Ramakrishnan R, Livny M (1996) BIRCH: an efficient data clustering method for very large databases. ACM SIGMOD Rec 25(2):103–114
20. Keogh E, Chu S, Hart D, Pazzani M (2001) An online algorithm for segmenting time series. In: International conference on data mining, pp 289–296
21. Bedi KP (2015) Clustering of categorized text data using cobweb algorithm. Int J Comput Sci Inf Technol Res 3(3):249–254
22. Kaneriya A, Shukla M (2015) A novel approach for clustering data streams using granularity technique. In: 2015 international conference on advances in computer engineering and applications, pp 586–590. https://doi.org/10.1109/ICACEA.2015.7164759
23. Yui M, Kojima I (2013) A database-hadoop hybrid approach to scalable machine learning. In: IEEE international congress on big data, pp 1–8

24. De Morales G (2013) SAMOA: a platform for mining big data streams. In: Proceedings of the twenty second international conference on world wide web, pp 777–778
25. Tamboli J, Shukla M (2016) A survey of outlier detection algorithms for data streams. In: 2016 3rd international conference on computing for sustainable global development (INDIACom), pp 3535–3540
26. Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E (2015) Deep learning applications and challenges in big data analytics. J Big Data 2(1):1
27. Jordan MI, Mitchell TM (2015) Machine learning with trends, perspectives, and prospects. Science 349:255–260
28. Keshvani T, Shukla M (2018) A comparative study on data stream clustering algorithms. In: International conference on computer networks, big data and IoT. Springer, Cham
29. Shukla M, Kosta YP (2016) Empirical analysis and improvement of density based clustering algorithm in data streams. In: 2016 international conference on inventive computation technologies (ICICT), pp 1–4. https://doi.org/10.1109/INVENTIVE.2016.7823262
30. Shukla M, Kosta YP, Jayswal M (2017) A modified approach of OPTICS algorithm for data streams. Eng Technol Appl Sci Res 7(2)
31. Amini A, Saboohi TYH (2014) On density-based data streams clustering algorithms: a survey. J Comput Sci Technol 29(1):116–141
32. Gomes Bezerra C, Costa BSJ, Guedes LA, Angelov PP (2019) An evolving approach to data streams clustering based on typicality and eccentricity data analytics. Inf Sci. https://doi.org/10.1016/j.ins.2019.12.022

Social Commerce Platform for Food Enthusiasts with Integrated Recipe Recommendation Using Machine Learning

Atul Kumar, Bhavak Kotak, Umang Pachaury, Preet Patel, and Parth Shah

Abstract In the contemporary world, humans are constantly involved in the race of making their lifestyle better by making all possible efforts to earn more and more money. However, it can be frequently observed that in this hustle for money, he often compromises on his health, and eating on the go has now become a trend that has replaced the infamous cook by your own tradition. Cooking at home by ourselves ensures the best quality of food and that ensures good health, whereas the eat on the go culture has led to more and more consumption of fast food which affects the health of a person thereby indirectly reducing his/her efficiency to work by deteriorating his/her health. Through this paper, we aim to help people around the globe with the problem stated above via a social commerce platform.

Keywords Social commerce · Recipe recommendation system · Machine learning

1 Introduction

Due to the pursuit of making our lives better, people often make compromises on their health. This is evidenced by the increasing number of people who eat on the go. The importance of cooking at home goes beyond good food quality and good health. It is also beneficial for the body as it helps maintain its optimum health. We hereby address this issue by introducing a social commerce platform where the users would find healthy recipes and would also be able to order the ingredients in the same proportion which would be described in the recipes. They would also be able to connect to nearby foodies as well as upload their recipes and cooking ideas/tutorials, etc. Aside from that, we will create a recipe recommendation model that will suggest food or recipes to consumers based on their moods and mouth feel.

A. Kumar (✉) · B. Kotak · U. Pachaury · P. Patel · P. Shah
Parul University, Vadodara, India
e-mail: atul.kumar8844@paruluniversity.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_49

2 Related Works

As seen in [1], the authors proposed that hybrid filtering (which combined collaborative and content-based filtering) the writers here made certain assumptions, such as considering all food ingredients in a recipe to be equally weighted and ignoring the weights of specific food items, their cooking methods, and so on, which resulted in great accuracy? Combination effects, etc., and transferred ratings from recipes to food and vice versa.

In [2], the creators proposed a chart-based calculation for menu arranging. The proposed method includes creating a formula chart to capture the co-event relationships between menu plans from various menus. On the developed formula chart, the approximated Steiner tree algorithm generated a menu. The proposed strategy achieves empowering effects, just as the character of mandated plan arrangements was serious to the initial ones, according to a review of menu selections. The authors offered two computations to advise elective fixes in their work Recommendation System for Alternative-Ingredients Based on Co-Occurrence Relation on Recipe Database and the Ingredient Category. Every one of the proposed recipes passed the cooking and tasting tests for each planned reason, a strategy was plausible.

In [3], the authors suggested a substitute ingredient recommendation technique based on a co-occurrence relationship in a recipe database. Two algorithms are suggested in this research to suggest alternate ingredients. Cooking and tasting tests were conducted. It was demonstrated that every one of the proposed techniques was productive for each arranged reason. Through cooking showing with the suggested elective fixing and abstract assessment tests, it was guaranteed that the two calculations suggested adequate elements for more than 90%. Be that as it may, Calculation 1 probably will not prescribe elective fixing like the trade fixing.

In [4], the authors proposed a dish recommendation system that took commonly used phrases as input and the system which they came up with displaying a comprehensive list of dishes along with their exact recipes as output which consisted of the precise varieties of supplements to take care of one's health. To do so, they first constructed a co-occurrence database, which identified the co-occurrence of 45 common nutrients with nouns like cold, acne, and so on, and typical nutrients including proteins, vitamins, fats. Then, they created an ingredient nutritional database with over 1800 well-known substances. Finally, they created a recipe database from 800,000 recipes found on www.cookpad.com. They figured out each recipe to know what ingredients and what amount of those ingredients were used and utilized that information to determine the in a dish, the amount of a nutrient. After then, the results were integrated with the co-occurrence database, so that the best suitable option which could treat the user's medical condition could be selected. They started by creating a system that could be utilized with an iPhone. So that, it would be more convenient as the user could hold the phone in one hand and refer to the recipe whenever necessary. They analyzed this system manually and compared it with the system developed for the iPhone based on the analysis of a thousand recipes from numerous sources. They observed that their system could recommend a meal

that includes items that are high in nutrients to help with health problems. The user's previous search history was ignored by their system, so it provided identical results if the user searched for the same item every day. Also, they wanted to further research on ways to avoid recommending such recipes to users who had ingredients to which the user was allergic.

In [5], the authors have tried to build a recipe recommendation model using machine learning. They have mainly focused on Indian Cuisine. Their project also consists of food pairing and alternative ingredient suggestions. To collect the recipes/data, they used web scraping. They used web scraping on Yummly.com for retrieving the data and analyzing it. For alternative ingredient recommendations, they used the Word2Vec model to get the nearest alternative of the ingredient. For food pairing, they used TF-IDF and cosine similarity to see which ingredients can be paired together. The drawback of this project is that they do not display any actual conclusions or figures to support the model's accuracy.

In [6], the creator investigates two principal partners in online informal organizations: the OSN clients and the OSN itself. Clients concede a great measure of actual recognizable data on OSNs, including physical, mental, social, and particular ascribes. Because of the assortment and explicitness of the individual data shared on OSNs, clients open themselves to danger for various digital just as actual assaults; for example, use third-party social media apps, as well as photographs on publicists' adverts. These compounds may cause data leakage from clients. Furthermore, clients' information gathered from various interpersonal groups causes linkage challenges; nonetheless, by linking the client via informal communities, a remarkably broad picture of the client can be constructed. The fear of being protected is grouped by the maker. And, security issues into the security and security issues classes in OSNs are given below. Data and content spillages and connections among clients on the OSN, interventions Sybil attacks on clients who have leaked into other clients. Accounts that have been compromised due to leaks into social media applications social spam and malware leaks into the OSN aggregator links distributed denial of service (DDoS). The unwelcomed openness of client data, combined with the OSNs-induced obscurity between the expert and specific aspects of the client's life, was situations of impeding results. Also, informal communities themselves purposely or accidentally appropriate anonymized social information utilized for de-anonymization, and induction assaults are adding to breaks in client privateness. Additionally, the tremendous volume of individual data, either uncovered by the normal client denied of adequate innovative information or because of OSNs' inability to bear the cost of complex protection apparatuses, has enraptured various associations that gather and sell clients interpersonal organization information.

In [7], the authors focused on the recommendation systems of recipes but focused on nutritional-based diets and to meet user's personalized dietary needs recommendations. They used two algorithms *k*-means clustering algorithm and a collaborative filtering algorithm to find the top nutritious meals from different cuisines and to suggest different healthy and balanced recipes. They took 40 students as an object for dietary needs of the past month for analysis. They took randomly any ten consecutive

days data, then they counted the number of foods that was recommended successfully. They prepared a user dashboard for the same project for the end-users.

In [8], nowadays, a big extent of facts is being accumulated to attain a higher decision-making manner. Many exclusive sources of facts, inclusive of social community interactions or servers, are constantly monitoring what Internet users are doing, and the data is gathered and categorized, despite the fact that the data is utterly heterogeneous. Since practically the discovery of the generation of the record within the early nineteenth century, our capability for data technology has never been so effective and significant. As a large amount of unprocessed data is accumulated, extracting knowledge from large datasets is becoming increasingly difficult. A fresh and extremely effective facts illustration is proposed for efficiently managing several heterogeneous datasets. In this way, it simplifies and reorganizes data gadgets in order to shorten data length and provide faster access to recorded records, while guaranteeing the validity, accuracy, usability, and integrity of the unique facts. For the give-up user, the offered facts shape is completely hidden and obvious. It is not necessary to have any prior knowledge of how the facts are compiled. An analogous API is also offered, which is clean and simple but expressive.

In [9], the authors suggested that models can be created using machine learning approaches. To finish a formula what the machine learning model proposed is non-negative network factorization. This methodology gives a straight non-negative part-based depiction of information. Be that as it may, it overlooks the data about flavor factors, and subsequently, the outcomes are wrong.

In this paper [10], creators confidently proposed an effective technique that suggests specific recipes of Indian cooking based on accessible fixings and traditionally enjoyed the delicious food. For this work, creators webbed scratching to make an assortment of plan assortment, and after that, they utilize the substance-based methodology of machine learning to suggest the plans. This framework offers the suggestions of Indian Cuisines dependent on fixings. Authors performed web scraping to collect data from various sites and gathered data in their local database in table format. And after properly cleaning of data, they intentionally used the bags-of-words method. Because, their recommendation model recommends recipes (food items) typically based on key components of a particular recipe. As a result, they selected the component column of their dataset and formed bags-of-words for every recipe. Bags-of-words consists of the paternoster for each recipe, and based on the analogy of those paternosters, they rank other food dishes in descending order of their analogy. In the recommendation part, they took the food item name as input and gave outputs as identical (matching) dishes. Initially, they provide the index of the particular food item which is inputted by the user and then they construct a sequence with a resemblance score with the help of a cosine resemblance matrix. At that time, they followed the pointer of the first two most related recipes and suggested those dishes to the user.

In [11], as evaluations, creators have used two approaches for recommending plans based on client preferences. Even though the thing-based methodology had a superior running season, the client-based methodology had more appropriate offers.

With 80% prepared information and 20% test information, the client-based cooperative separation alongside the neighborhood of 200 clients with Euclidean distance similitude would yield more results (Tables 1, 2, 3, and 4).

Table 1 Ninety percent of the data was used in the training process, whereas ten percent was used in the testing process

Similarity	$n = 100$	$n = 200$	$n = 300$	$n = 500$	$n = 1000$
Euclidean distance (AAD)	0.7745	0.7484	0.7511	0.7667	0.820
Person correlation (AAD)	0.83	0.7906	0.7782	0.8133	0.842
Person correlation (RMSE)	0.9783	0.9593	0.9641	0.9697	1.000
Euclidean distance (RMSE)	1.04	0.9993	1.006	1.005	1.065

Table taken from [11]

Table 2 Data was utilized for training in 80% of the cases and testing in 20% of the cases

Similarity	$t = 0.9$	$t = 0.8$	$t = 0.7$	$t = 0.6$	$t = 0.5$
Person correlation (AAD)	0.9140	0.8771	0.8624	0.8339	0.8207
Euclidean distance (AAD)	0.8795	0.8919	0.8714	0.8022	0.7488
Euclidean distance (RMSE)	1.1097	1.2740	1.1037	1.0090	0.9555
Person correlation (RMSE)	1.1177	1.009	1.1005	1.0486	1.0424

Table taken from [11]

Table 3 Ninety percent of the data was used for training, with the remaining ten percent for testing

Similarity	$n = 2$	$n = 4$	$n = 6$	$n = 8$	$n = 16$
Person correlation (AAD)	1.663	1.69	1.69	1.576	1.639
Euclidean distance (AAD)	3.07	1.29	1.346	1.166	1.18
Person correlation (RMSE)	2.1357	1.839	1.678	2.078	1.8231
Euclidean distance (RMSE)	2.07	1.265	1.801	1.693	1.5893

Table taken from [11]

Table 4 Using 80% of the training data and 20% of the test data, the following findings were achieved

Similarity	$t = 0.9$	$t = 0.8$	$t = 0.7$	$t = 0.6$	$t = 0.5$
Person correlation (AAD)	1.211	0.8977	1.6216	1.1686	0.9553
Euclidean distance (AAD)	1.666	0.8886	1.7776	1.95	1.2538
Person correlation (RMSE)	1.1956	2.2223	1.9247	1.8125	1.3195
Euclidean distance (RMSE)	2.84	2.0955	1.4513	2.2580	1.7751

Table taken from [11]

Machine learning-based food recipe recommendation system exact suggestions for all recipe datasets. A restriction of 0.8 with Euclidean distance similitude for the reproduced dataset would provide us with a large number of proposals. When compared to a thing-based shared filtering strategy, the client-based methodology was shown to be more suitable and performed better. Because the quantity of client-to-client and thing-to-thing partnerships is large in the all recipe dataset, we find the client-based methodology to be superior to the reproduced dataset. The client-based methodology's aftereffects are found to be superior to the thing-based methodologies. This is a summarized way for recommending plans using client-based methodology. The equivalent can be combined with a content-based approach to deal with acquiring a mixture approach, and it can be used in the clinical field to recommend medications for patients suffering from various illnesses, as well as in the pharmaceutical industry to recommend various medications for various infections.

In [12] the culinary world, a dish can be made with different approaches, and the flavor depends on how you use certain ingredients or what approach or mechanism you take to make a recipe. All of this can elevate a dish from street style vendor to five stars. All of this requires a lot of extensive research and work, and RecipeScape solves this. The authors have presented an interactive tool to browse and compare many recipes of a single dish. They give a tree representation of recipes. They also present a computational pipeline that extracts cooking processes from recipe text and calculates a procedural similarity between them. During their research, they identified three design goals for their system and made three tools addressing each design goal:

1. Design Goal 1 (De-1): Provide statistical data on ingredients and cooking procedures.
2. Design Goal 2 (De-2): Individual recipes should be examined and compared in detail.
3. Design Goal 3 (De-3): In order to provide a comprehensive study of recipes, you must do it in bulk.

Based on the design goals, they made RecipeScape which had three important components:

- RecipeMap: Provides a summary of recipes together with cluster information. Taking care of the Design Goal 3.
- RecipeMap: Groups of recipes based on structural similarity.
- RecipeDesk: Provides information on individual recipes, including the original description, the matching tree representation, and pairwise recipe comparison. This is how Design Goal 2 is addressed.

RecipeStat: This shows the patterns of cooking activities and ingredients used in RecipeMap's selected clusters of recipes. This relates to the first design goal. To translate recipe text into tree representation and calculate the pairwise distance between recipes, they employ a computational pipeline that includes part-to-speech tagging and human annotation.

3 Conclusion

The food industry is expanding continuously and researching food to offer unique taste, new vision, and a new experience to the client, but in the search for dignity, they sometimes fail to remember the obvious and the valuable step of customer satisfaction. Eating out regularly in eateries deprives the feel of homely food, the process of cooking your meal, and lacking the enjoyment of the fruit of one's cooking endeavors. With the examined model of machine learning, the recipe recommendations to the cravings of the user for what the user desires to eat providing the terms like the taste, preparation time, nutrition, healthy, without any compromise to their choices in their hands, prepared by them, appreciated by them at home preserving the homely feel, hygiene, and most efficient everything delivered at their footstep with all the fresh ingredients as per the recipe they wanted. The presented model employs both the collaborative filtering process and the *k*-means clustering technique. The most difficult component of this generative model was mass evaluation, but we overcome it by understanding the output with *k*-means clustering, which we then filtered with a collaborative filtering strategy. According to our findings, the MIL structure, a decreased vocabulary size of 150,000, a greater number of hidden cells at 512, and a shorter sequence length of 25 all aid in the development of recipes with the desired recipe structure are close to existing recipe clusters and make linguistic sense. Normal assignments incorporate arranging various ways to deal with cooking a dish and recognizing utilization examples of specific fixings or cooking strategies, all of which require broad perusing and examination of numerous plans. Nonetheless, no current framework was offering help for such top to bottom and at-scale examinations. So, we utilized RecipeScape, which bunches plans into unmistakable methodologies and catches remarkable use examples of fixings and cooking activities. We also used recipe recommendations based on user preferences that the results were near to 60–80%, and we targeted to make it at least 90% by combining the models with the new algorithms and proper filtration. We are also aiming for healthier recipes for the patients who are suffering from any kind of disease to have proper nutrition-rich meals without any compromises. We also proposed alternative food ingredients models to solve the problem of the expanding measure of electronically accessible information that has brought about progressively huge datasets, which surpass the limit of customary AI calculations concerning memory utilization and execution in a sensible time. The objective is not to propose new efficient calculations, however, to present another information structure that could be utilized by an assortment of existing calculations without altering their unique schemata.

References

1. Freyne J, Berkovsky S (2010) Intelligent food planning: personalized recipe recommendation. In: Proceedings of the 15th international conference on intelligent user interfaces, Feb 2010, pp 321–324

2. Kuo FF, Li CT, Shan MK, Lee SY (2012) Intelligent menu planning: recommending set of recipes by ingredients. In: Proceedings of the ACM multimedia 2012 workshop on multimedia for cooking and eating activities, Nov 2012, pp 1–6
3. Shino N, Yamanishi R, Fukumoto J (2016) Recommendation system for alternative-ingredients based on co-occurrence relation on recipe database and the ingredient category. In: 2016 5th IIAI international congress on advanced applied informatics (IIAI-AAI), July 2016. IEEE, pp 173–178
4. Ueta T, Iwakami M, Ito T (2011) A recipe recommendation system based on automatic nutrition information extraction. In: International conference on knowledge science, engineering and management, Dec 2011. Springer, Berlin, Heidelberg, pp 79–90
5. Maheshwari S, Chourey M (2019) Recipe recommendation system using machine learning models. Int Res J Eng Technol (IRJET) 6(9):366–369
6. Kayes I, Iamnitchi A (2017) Privacy and security in online social networks: a survey. Online Soc Netw Media 3:1–21
7. Yuan Z, Luo F (2019) Personalized diet recommendation based on K-means and collaborative filtering algorithm. J Phys Conf Ser 1213(3):032013. IOP Publishing
8. Padillo F, Luna JM, Cano A, Ventura S (2016) A data structure to speed-up machine learning algorithms on massive datasets. In: International conference on hybrid artificial intelligence systems, Apr 2016. Springer, Cham, pp 365–376
9. De Clercq M, Stock M, De Baets B, Waegeman W (2016) Data-driven recipe completion using machine learning methods. Trends Food Sci Technol 49:1–13
10. Nilesh N, Kumari M, Hazarika P, Raman V (2019) Recommendation of Indian cuisine recipes based on ingredients. In: 2019 IEEE 35th international conference on data engineering workshops (ICDEW), Apr 2019. IEEE, pp 96–99
11. Vivek MB, Manju N, Vijay MB (2018) Machine learning based food recipe recommendation system. In: Proceedings of international conference on cognition and recognition. Springer, Singapore, pp 11–19
12. Chang M, Guillain LV, Jung H, Hare VM, Kim J, Agrawala M (2018) Recipescape: an interactive tool for analyzing cooking instructions at scale. In: Proceedings of the 2018 CHI conference on human factors in computing systems, Apr 2018, pp 1–12
13. Saravanakumar K, Deepa K (2016) On privacy and security in social media—a comprehensive study. Procedia Comput Sci 78:114–119
14. Lewis T (2018) Digital food: from paddock to platform. Commun Res Pract 4(3):212–228
15. Ghahtarani A, Sheikhmohammady M, Rostami M (2020) The impact of social capital and social interaction on customers' purchase intention, considering knowledge sharing in social commerce context. J Innov Knowl 5(3):191–199
16. Funk S (2012) Social networks and the promotion of regional food culture
17. Zhou L, Zhang P, Zimmermann HD (2013) Social commerce research: an integrated view. Electron Commer Res Appl 12(2):61–68
18. Turban E, Bolloju N, Liang TP (2010) Social commerce: an e-commerce perspective. In: Proceedings of the 12th international conference on electronic commerce: roadmap for the future of electronic business, Aug 2010, pp 33–42
19. Ueda M, Takahata M, Nakajima S (2011) User's food preference extraction for personalized cooking recipe recommendation. In: Workshop of ISWC, Oct 2011, pp 98–105
20. Elsweiler D, Trattner C, Harvey M (2017) Exploiting food choice biases for healthier recipe recommendation. In: Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval, Aug 2017, pp 575–584
21. Starychfojtu J, Peska L (2020) SmartRecepies: towards cooking and food shopping integration via mobile recipes recommender system. In: Proceedings of the 22nd international conference on information integration and web-based applications & services, Nov 2020, pp 144–148
22. Zhang C, Sun J, Zhu X, Fang Y (2010) Privacy and security for online social networks: challenges and opportunities. IEEE Netw 24(4):13–18
23. Willis A, Lin E, Zhang B (2017) Forage: optimizing food use with machine learning generated recipes

24. Svistova SF (2018) Analyzing cooking recipes with machine learning. 8(74)
25. Forbes P, Zhu M (2011) Content-boosted matrix factorization for recommender systems: experiments with recipe recommendation. In: Proceedings of the fifth ACM conference on recommender systems, Oct 2011, pp 261–264
26. Sharma PP, Gupta A, Parmar H, Patidar D, Rana APK (2021) Food recipe recommendation system
27. Garrido-Merchán EC, Albarca-Molina A (2018) Suggesting cooking recipes through simulation and bayesian optimization. In: International conference on intelligent data engineering and automated learning, Nov 2018. Springer, Cham, pp 277–284
28. Bernardes D. Food recipe recommendation and the social filtering formalism

CoviFacts—A COVID-19 Fake News Detector Using Natural Language Processing

Nandita Kadam, Shreyas More, Vishant Mehta, Sanyam Savla, and Krisha Panchamia

Abstract Fake news confronts us on a daily basis in today's fast-paced social media world. While some instances of fake news might seem innocuous, there are many examples that prove to be menacing. Misinformation or disinformation which takes the form of these weaponized lies which eventually amount to defective information, defamatory allegations, and hoaxes. The only motive behind such a malicious act is to engender emotional instability among the public. One such prevalent example today is COVID-19 which has caused an unprecedented paradigm shift in numerous businesses and quotidian activities across the globe. One of the primary activities is being news reporting. On average, people are spending almost one hour a day reading news via many different sources. The development in technology has obviated the barriers between sharing of information, thereby truly making the industry cosmopolitan. Therefore, it is paramount to curb fake news at source and prevent it from spreading to a larger audience. This paper describes a system, where the user can identify apocryphal news related to COVID-19 so as to ensure its authenticity.

Keywords Fake news detection · Natural language processing · COVID-19 · Machine learning · News classification · Feature engineering · Logistic regression · SVM · AdaBoost classifier · MultinomialNB

N. Kadam (✉) · S. More · V. Mehta · S. Savla · K. Panchamia
K. J. Somaiya College of Engineering, Vidyavihar, Mumbai, India
e-mail: nandita.kadam@somaiya.edu

S. More
e-mail: shreyas.mm@somaiya.edu

V. Mehta
e-mail: vishant.m@somaiya.edu

S. Savla
e-mail: sanyam.savla@somaiya.edu

K. Panchamia
e-mail: krisha.panchamia@somaiya.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_50

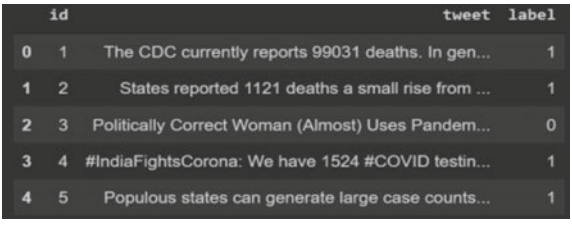
1 Introduction

Social media for news consumption is a deleterious platform. Despite being easy to access, it promulgates rapid dissemination of information that is low quality and fatal. It is a false narrative that is promoted to alter the beliefs and perceptions of people, ultimately leading to the infiltration of society. As the COVID-19 virus quickly spreads across the globe, the world is not just dealing with a "pandemic" but also an "infodemic". Considering that the world is on the brink of disaster, it is paramount to validate the authenticity of dubious information. Fake news detection on social media presents unique challenges that make the existing detection algorithms for traditional news media to be ineffective [1–4]. Thus, we present in this paper a web application whose detection process is based on the assemblage learning of machine learning classifiers and natural language processing (NLP) [5] on the collected dataset for COVID-19 misinformation, to bifurcate fake and real news at its maximum capacity. The paper begins with an introduction and is followed by methodology in Sect. 2. Section 3 explicates the proposed system. The experimental results are revealed in Sect. 4. The paper ends with a conclusion and potential future scope in the final section, Sect. 5.

2 Methodology and Proposed Solution

2.1 Data Gathering

The preliminary stage pertaining to the problem statement is to finalize a well-diverse and appropriate dataset. The machine learning models used in this system are trained on a COVID-19 Fake News Dataset (COVID-19 Fake News Detection in English) [6, 7]. This dataset is manually annotated and curated with 10,700 social media posts and articles of real and fake news on COVID-19. The dataset entails textual facts as well as URL links to different articles related to coronavirus. All these tweets are either labeled as "0" (fake) or "1" (real) (see Fig. 1). The dataset comprises three splits as follows:



	id	tweet	label
0	1	The CDC currently reports 99031 deaths. In gen...	1
1	2	States reported 1121 deaths a small rise from ...	1
2	3	Politically Correct Woman (Almost) Uses Pandem...	0
3	4	#IndiaFightsCorona: We have 1524 #COVID testin...	1
4	5	Populous states can generate large case counts...	1

Fig. 1 Example of real (1) and fake (0) news from the dataset

	id	tweet	label	total
0	1	The CDC currently reports 99031 deaths. In gen...	1	the cdc currently report 99031 death in gener...
1	2	States reported 1121 deaths a small rise from ...	1	states reported 1121 death small rise last tu...
2	3	Politically Correct Woman (Almost) Uses Pandem...	0	politically correct woman almost uses pandemi...
3	4	#IndiaFightsCorona: We have 1524 #COVID testin...	1	indiafightscorona we 1524 covid testing labor...
4	5	Populous states can generate large case counts...	1	populous state generate large case count look...

Fig. 2 First five rows of the dataset before and after applying data preprocessing techniques

– **Training split**: Consists of 6420 rows.
– **Validation split**: Consists of 2140 rows.
– **Test split**: Consists of 2140 rows.

2.2 Data Preprocessing

The following steps are followed as a part of data preprocessing prior to applying NLP techniques (see Fig. 2).

– **Handling of missing/null values** [5]: The dataset was checked for null entries using "imputation", which validated that the dataset was balanced and did not contain any null values.
– **Cleaning of the dataset**: Textual data cannot be used directly due to unwanted words or special symbols. Hence, "REGEX" was used to remove punctuation or any sub-string that does not hold words or spaces.
– **Tokenization**: All the sentences or phrases in the dataset were split into individual words using the "nltk python library".
– **Removal of stopwords**: To classify the text into different categories, more focus is given to the words defining the meaning or context of that text by removing the stop words using the "nltk.corpus" Python library.
– **Lemmatization** [5]: Converts the text (suffixes or prefixes) into its root form using the "nltk.stem" Python library. For instance: "learnt, learned, learning" gets converted into "learn" (base form).

2.3 Applying NLP Techniques

Tf-idf Vectorization [5]: Term frequency-inverse document frequency (tf-idf) is used to convert raw output texts into features to normalize unnecessary words. In plain words, it is the extraction of features from a text to convert it into a vector of numbers. These features are converted into an array in which the weights are assigned

to words in the dataset, based on their frequency of occurrence. This vectorization is implemented using the “tf-idf vectorization” in “scikit learn”.

2.4 Input

– Inputs to the system are accepted in the form of textual summary or URL of news article/tweet.
– News articles that are solely related to the pandemic will be accepted for producing the predictive text.
– Other domains of articles will not produce any predictive text.

2.5 Output

– The text or `news_url` will either produce a predictive text of “real” or “fake”.
– Alongside this, it will also specify the author, title, keywords, and short summary of the `news_url` given as an input.
– In case of any errors in the input validation, error messages will be displayed on the UI accordingly.

3 Approach to the Proposed System

3.1 System Architecture

The process starts with data acquisition, followed by data cleaning and exploration, extraction of linguistic features, and feature extraction. The modeling and training methods subsume a comparison between four machine learning classifiers, i.e., AdaBoost classifier [8], logistic regression [9], support vector machine (SVM) [8], and MultinomialNB [10]. The above-mentioned algorithms were considered for accuracy, and the best-fitted model was chosen for the COVID-19 fake news classification problem statement. Figure 3 is the true representation of the system architecture.

3.2 Web Scraping and Pipelining

– In the pursuit of predicting the URL links of any COVID-19-related article, data scraping from the web was of paramount importance.

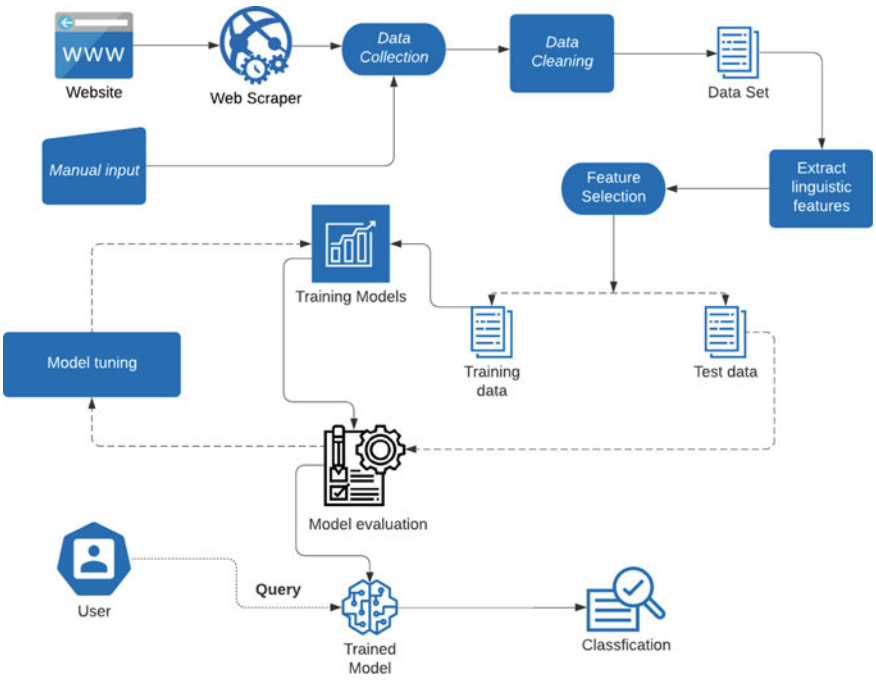


Fig. 3 System architecture of CoviFacts

– Data was scraped and curated from the URL link that was provided by the user as an input.
– Web scraping was implemented in this application via the "Newspaper3k python3 library". "Newspaper" has a seamless language extraction and detection. If no language is specified, "Newspaper" will attempt to auto-detect a language.
– The machine learning model was first serialized into a single pickle file, i.e., "model.pickle" using the process of pipelining for ease of further prediction of the data.
– Using the inbuilt functions, the summary of the article was scraped and passed as a parameter whose likelihood is to be predicted by the chosen model.

4 Experimental Results

The proposed system was implemented by using four machine learning algorithms during the modeling and training stage. The algorithms were trained on the training dataset and were further evaluated on the basis of their performance on the test dataset comprising 2140 samples.

Table 1 Comparison of ML classifiers

Algorithms	Confusion matrix metrics				Accuracy (test set) (%)
	TN	FP	FN	TP	
AdaBoost	890	200	164	886	83
Logistic regression	980	100	73	987	92
SVM	1063	51	60	966	95
MultinomialNB	923	153	82	982	89

Fig. 4 Evaluation metrics for SVM

	precision	recall	f1-score
fake	0.95	0.93	0.94
real	0.94	0.96	0.95
accuracy			0.95
macro avg	0.95	0.95	0.95
weighted avg	0.95	0.95	0.95

AdaBoost Classifier [8] is an adaptive boosting statistical classification meta-algorithm with a training accuracy of 91%. **Logistic regression** [9] is used for binary classification, based on the logistic function and sigmoid function and provides a training accuracy of 100%. **SVM** [8] uses term-frequency in order to estimate the maximum likelihood based on the training data and has an accuracy of 100% on the same. **MultinomialNB** [10] is suitable for classification with discrete features (word count/textual data) and leverages a training accuracy of 96. The dataset is split using an 80:20 ratio, and the percentage accuracy and confusion matrix were used as the evaluating metrics for the classification model as described in Table 1.

As clearly evident from Table 1 and Fig. 4, support vector machine (SVM) has leveraged maximum accuracy, precision, and $F1$-score for the given dataset by providing the least generalization error on the test set. With an accuracy of **95%** within the fake news classification system, **SVM** [8] was chosen as the classification algorithm for this project.

The experimental results obtained on the web browser for various articles are illustrated below (see Figs. 5 and 6):

1. **Textual Data Prediction**
   Input Given (test set):
   Low #vitaminD was an independent predictor of worse prognosis in patients with COVID-19.
   Expected Output*: Real (1)*
   Predicted Output*: Real [True Positive]*

Fig. 5 True positive in textual input

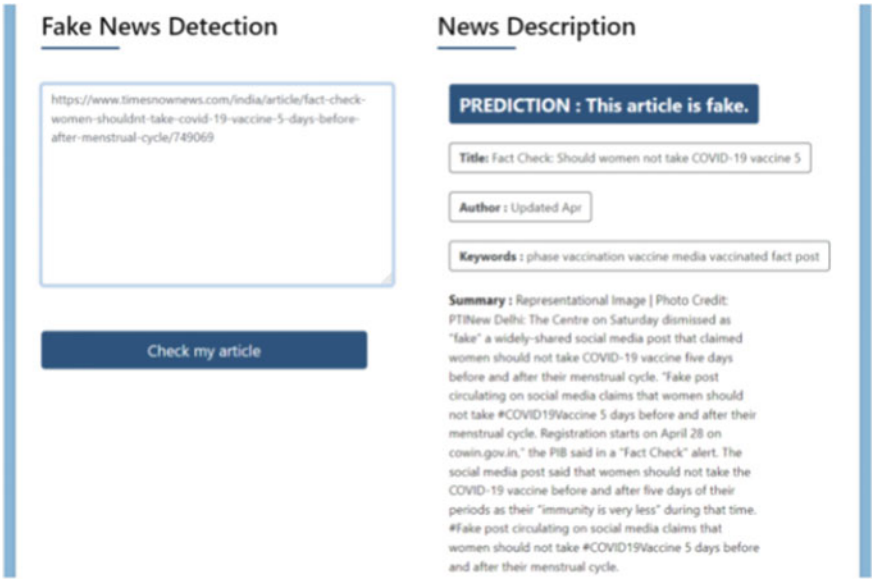


Fig. 6 True negative in URL input with key descriptions like title, author, summary, and keywords

2. **URL Link Prediction**
   Input Given:
   https://www.timesnownews.com/india/article/face-check-women-shouldnt-take-covid-19-vaccine-5-days-before-after-menstrual-cycle/749069
   Expected Output*: Fake (0)*
   Predicted Output*: Fake [True Negative]*

5 Conclusion and Future Scope

The task of classifying fake and real news requires skill and copious amounts of domain expertise to perform efficiently. This web-based application ("CoviFacts") provides comprehensive functionalities to predict the authenticity of a news article in textual as well as URL format and has reached a system accuracy of over 95%.

The primary aim of the research is to identify patterns in text that differentiate fake articles from true news using machine learning models, ensemble techniques, and NLP. The prescient impetus is to provide the general public easy access to analyze a skeptical piece of information in these times of rumors and misleading information.

It is often said that fake news detection is a very critical problem statement, and no 100% accuracy can be obtained in this domain. Therefore, future work entails the formulation of better algorithms and optimizers to increase the accuracy to at least 99%.

The future aim is to perpetuate developing this application with the motivation to hypertune the model and increase its accuracy with each iteration. In terms of the project predictions, the future scope incorporates developing a Google Chrome extension that would send a prompt while hovering over any COVID-19 news and will direct the user to the detection page for checking the validity of the article they are reading.

References

1. Hiramath CK, Deshpande GC (2019) Fake news detection using deep learning techniques. In: 2019 1st international conference on advances in information technology. https://doi.org/10.1109/ICAIT47043.2019.8987258
2. Poddar K, Geraldine Bessie Amali D, Umadevi KS (2019) Comparison of various machine learning models for accurate detection of fake news. In: 2019 innovations in power and advanced computing technologies (i-PACT). https://doi.org/10.1109/i-PACT44901.2019.8960044
3. Zhang S, Wang Y, Tan C (2018) Research on text classification for identifying fake news. In: 2018 international conference on security, pattern analysis, and cybernetics (SPAC). https://doi.org/10.1109/SPAC46244.2018.8965536
4. Kousika N, Deepa S, Deephika C, Dhatchaiyine BM, Amrutha J (2021) A system for fake news detection by using supervised learning model for social media contents. In: 2021 5th

international conference on intelligent computing and control systems (ICICCS). https://doi.org/10.1109/ICICCS51141.2021.9432096
5. Bharadwaj A, Ashar B, Barbhaya P, Bhatia R, Shaikh Z (2020) Source based fake news classification using machine learning. Int J Innov Res Sci Eng Technol (IJIRSET) 9(6)
6. Pathwa P, Sharma S, Gupta V, Kumari G. Fighting an infodemic: COVID-19 fake news dataset. Available at: https://arxiv.org/abs/2011.03327
7. Dataset. Constraint@AAAI2021—COVID19 fake news detection in English. Available at: https://competitions.codalab.org/competitions/26655#participate-get-data. Accessed 26 Sept 2021
8. Zhang Y, Ni M, Zhang C, Liang S, Fang S, Li R, Tan Z (2019) Research and application of AdaBoost algorithm based on SVM. In: 2019 IEEE 8th joint international information technology and artificial intelligence conference (ITAIC). https://doi.org/10.1109/ITAIC.2019.8785556
9. Abdullah-All-Tanvir, Mahir EM, Akhter S, Huq MR (2019) Detecting fake news using machine learning and deep learning algorithms. In: 2019 7th international conference on smart computing & communications (ICSCC). https://doi.org/10.1109/ICSCC.2019.8843612
10. Mandical RR, Mamatha N, Shivakumar N, Monica R, Krishna AN (2020) Identification of fake news using machine learning. In: 2020 IEEE international conference on electronics, computing and communication technologies (CONNECT). https://doi.org/10.1109/CONECCT50063.2020.9198610

Improved LBP Face Recognition Using Image Processing Techniques

G. Padmashree and A. K. Karunakar

Abstract The face recognition process is used to distinguish individual's faces based on their unique facial traits. In face recognition, the detection of faces in real-time videos under varying illumination conditions is one of the challenging tasks. In this study, we are detecting faces using Haar classifiers because of their high detection accuracy and local binary pattern (LBP) classifiers due to their invariant nature under varying illumination conditions. Image processing techniques such as contrast adjustment, bilateral filtering, histogram equalization, image blending, and quantization are applied to improve the detected faces. Also, we have applied quantization on raw face images at various levels to evaluate the feasibility of the proposed method in effectively recognizing the faces in low-quality images. Using local binary pattern histogram (LBPH) recognizer, a face recognition rate of 100% has been achieved when resized raw images and preprocessed images are blended. Also, an equal performance has been achieved when the quality of the images is reduced by applying quantization of 16 levels. Hence, the proposed method has proven its effectiveness in recognizing the faces in low-quality images. The results show that using the preprocessed image, the proposed face recognition method is invariant to varying illumination conditions.

Keywords Face recognition · Local binary pattern · Quantization · Face detection · Haar classifiers

1 Introduction

Face recognition is the concern of distinguishing or authorizing appearances. The process of face recognition involves detection of face, extraction of features from detected faces, and comparison with the faces in the database. Face detection methods are categorized [1] into four categories, knowledge-based, feature-based, template matching, and appearance-based methods. The commonly used face detection

G. Padmashree · A. K. Karunakar (✉)
Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India
e-mail: karunakar.ak@manipal.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_51

method is feature-based, where features are extracted based on intensity data without any information of the structure of the face [2]. Feature extraction aims at extracting pertinent features from the image and representing them in a lower dimensionality space. Finally, these features are compared with the features of the images stored in the database for recognition. Some of the challenges involved in face recognition are variations in illumination, occlusions, low-quality images, facial expressions, and so on. In our study, we focus on issues like varying illumination conditions and facial expressions using Haar and local binary patterns classifiers for facial feature extraction. Alongside, we perform face recognition on low-quality images by applying quantization at different levels on face images. To perform this study, we have created a dataset of face images of varying illumination and expressions.

2 Related Work

Ojala et al. [3] proposed a method to describe the local spatial structure of an image. Similarities between the textures of the images are evaluated using pattern spectrum, and histograms are compared. Kasinski and Schmidt [4] proposed the architecture and performance of the face and eyes based on the Haar classifier and achieved an accuracy of 94%. Chintalapati et al. [5] proposed a face recognition system for automated attendance management using Haar classifiers for face detection. Later, histogram equalization was performed on detected faces which were further processed for feature extraction using PCA and LBPH [6] algorithms.

For recognition, support vector machine(SVM), distance classifier, and Bayesian methods were applied, and they had obtained a recognition rate of 95% with local binary pattern histogram (LBPH) outperforming the others, and recognition of faces with variations in poses was a major drawback. Sharma et al. [7] suggested a hybrid strategy for pose invariant face recognition that combines three algorithms using PCA. Using Haar classifiers, we first detect the face and its part. The second step is to extract features from detected faces using LBP classifiers and finally to apply PCA on each extracted feature for recognition. Using this approach, 98% accuracy was achieved on images with varying facial expressions and poses. Using Haar classifiers, a smart attendance capturing and management system was proposed by Surekha et al. [8] under controlled and uncontrolled environments. 100% accuracy was obtained under a controlled environment. However, only they were able to achieve only 60% accuracy when tested under an uncontrolled environment.

A face recognition system was developed by Mady et al. [9] based on video transmission under varying lighting conditions, facial expressions, pose, alignment, occlusion, race variation, and gesture. In an uncontrolled environment, Haar classifiers were employed to detect faces. For face features extraction, histograms of oriented gradients and local binary pattern were combined, and for face classification, a random forest classifier (RF) was used. The recognition accuracy of 97.6% was obtained by using histogram equalization as one of the preprocessing steps.

From these research works, it is observed that face recognition was performed on high-quality images. However, the low-quality images with suppressed facial features degrade the performance of the face recognition system.

Furthermore, the performance of the existing face recognition system under varying illumination conditions, facial expressions, and poses was satisfactory. Thus, to address these issues, we enhance the quality of the images by applying image processing techniques such as contrast adjustment, bilateral filtering, histogram equalization, and image blending on the detected faces and perform recognition using an LBPH recognizer.

3 Methodology

In the proposed approach, as shown in Fig. 1, faces are detected using Haar [10] and LBP classifiers [6] in real-time videos captured using webcam. Then, to enhance the quality of detected face images, preprocessing such as contrast adjustment, filtering, histogram equalization, image blending, and quantization operations are performed on each detected face image. Using LBPH face recognizer, features from the preprocessed face images are extracted, and separate models are generated for each of the image categories as mentioned in Table 1.

3.1 Dataset

We created our dataset of face images in real-time videos captured using webcam. The subject was seated at a distance ranging between 3 and 6 feet in the line of sight of the

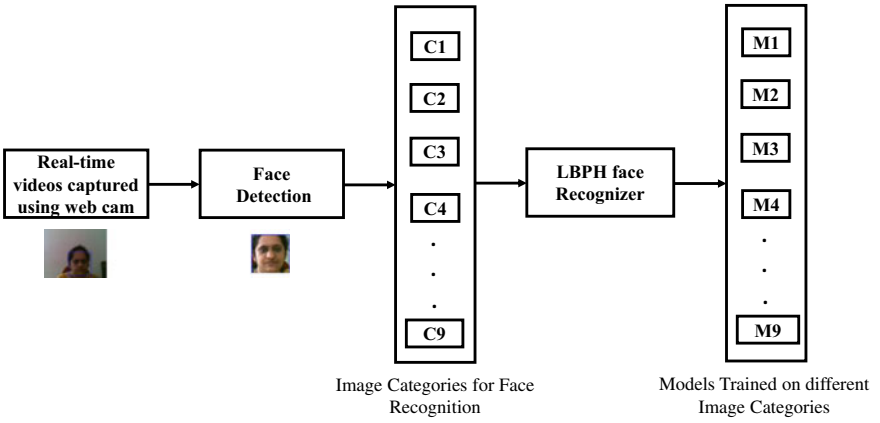


Fig. 1 Proposed method of face recognition

Table 1 Image category details

Category (C)	Description
C1	Raw images
C2	Histogram equalized images
C3	Blended images
C4	Quantized raw images (8 levels)
C5	Quantized raw images (16 levels)
C6	Quantized histogram equalized images (8 levels)
C7	Quantized histogram equalized images (16 levels)
C8	Quantized blended images (8 Levels)
C9	Quantized blended images (16 Levels)

webcam. The subject was monitored for 2 min, and frames were extracted at 2 frames per second (FPS). Using Haar and LBP classifiers, faces were detected from these frames and stored in JPEG format. The detected faces were of varying expressions and captured at different illumination conditions. Figure 2 shows the sample dataset of faces detected using LBP classifiers. For this study, we have considered six different subjects, and each subject is assigned a unique ID or name such that the faces can be recognized and provide precise output. Each subject has contributed 400 face images, and in total, the dataset comprised of 2400 images, out of which for training and testing, we used 300 and 100 face images, respectively, of each subject. The dataset was categorized into two groups, Group 1 consists of raw images of varying resolutions ranging between 79×79 and 190×190, and Group 2 consists of images resized to 128×128.

3.2 Face Detection

Classifiers are responsible for deciding whether a given image is face or non-face. To learn the process of classifying an image as face or non-face, numerous examples are used for training the classifier. The two types of classifiers used in face detection are Haar classifiers and LBP classifiers.

Haar classifiers Haar classifiers [10] were the first real-time face detectors used in object recognition. The algorithm works in four stages: Haar features selection, generation of integral image, AdaBoost training, and cascading classifiers. Some of the types of features detected by Haar classifiers, i.e., edge features and line features are depicted in Fig. 3. Convolution kernels, which yield a single result by subtracting the sum of pixels under the white rectangle from the sum of pixels under the black rectangle, are similar to Haar classifiers. Identifying these features in large images is a difficult task. Using an integral image, an intermediate representation of the image, features can be evaluated at a faster rate which results in identifying a large

Fig. 2 Sample faces detected using LBP classifiers

number of features. Among these features, many features might be irrelevant, and by using AdaBoost training, we selected the best features which classify the face and non-faces on the given set of images. A cascade of classifiers is used to reduce the computation time in recognizing the faces and non-faces in a given image.

Local Binary Pattern(LBP) Classifiers Local Binary Pattern(LBP) Classifiers [6] is one of the texture-based operators which has shown very high performance because of its invariance of lighting and robustness. The basic principle of LBP is to divide the input image into non-overlapping areas known as cells. The classifier then compares each of the core pixel's eight neighbors. If the central pixel's value is greater than or equal to the neighboring pixel's value, the neighboring pixel is set to 1, otherwise it is set to 0. The modified pixel values are then read clockwise, yielding a binary number, which is then translated to a decimal number, yielding the new value of the center pixel. Once the LBP value is evaluated, the histogram is generated. The complete process of the LBP histogram is shown in Fig. 4.

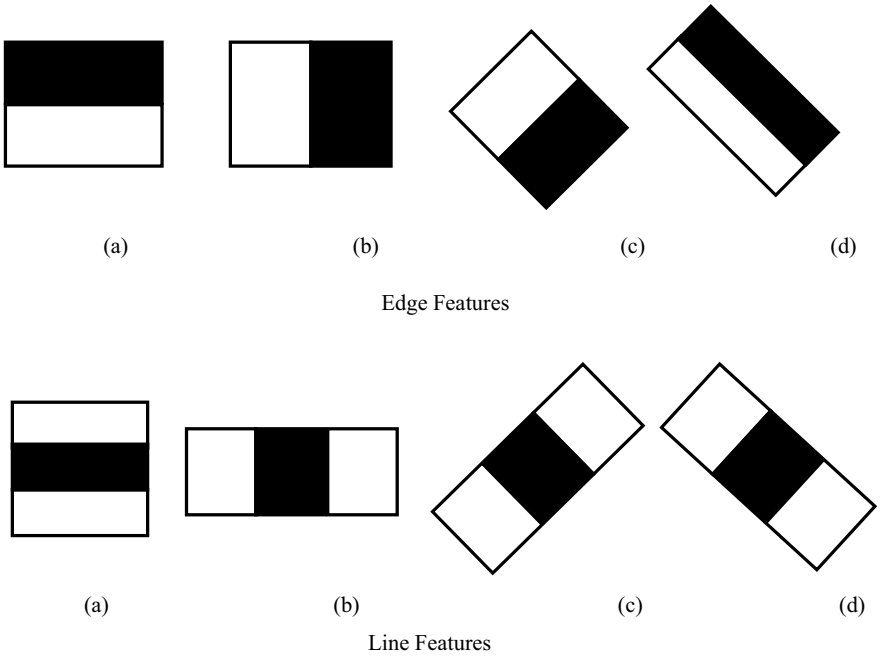


Fig. 3 Types of features of Haar classifiers

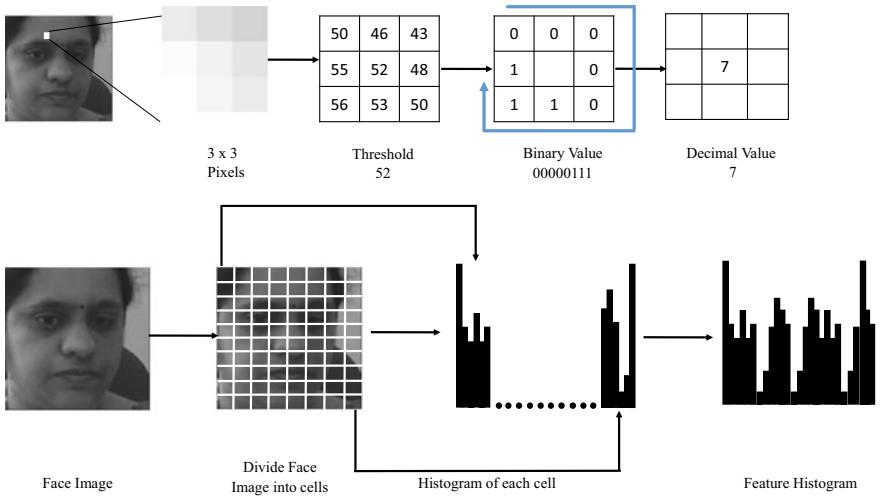


Fig. 4 General process of LBP operation

3.3 Image Categories

For recognition of faces, faces are detected using Haar classifiers and LBP classifiers. To improve the quality of the detected face, we apply contrast adjustment, bilateral filtering, histogram equalization, and image blending. Also, the quality of the images were degraded by applying quantization of 8 and 16 levels. The different category of images used in the proposed method are as shown in Table 1. In *Category 1(C1)*, raw images of faces detected using Haar classifiers and LBP classifiers are considered. In *Category 2(C2)*, image processing techniques such as contrast adjustment, bilateral filtering, and histogram equalization are applied on the raw images. In contrast adjustment, we improve the contrast of the detected faces [11], by remapping the pixel values to a new range of intensity values as defined in Eq. 1. Here, x is the target image, f is the source image, i and j are the pixel position in the source image. α is called the gain parameter, and β is called the bias parameter which are responsible for controlling the contrast and brightness of the given image. Experimentally, it was found that for the α value of 1.5 and β value of 0, better results were obtained.

$$x(i, j) = \alpha * f(i, j) + \beta \tag{1}$$

Then, bilateral filtering [12] is applied on these contrast enhanced images for smoothening the edges and removing noise in the images which is defined in Eq. 2, where normalization factor W_p confirms pixel weights sum to 1.

$$g(i, j) = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s}(\| p - q \|) G_{\sigma_r}(|x_p - x_q|) x_q \tag{2}$$

$$W_p = \sum_{q \in S} G_{\sigma_s}(\| p - q \|) G_{\sigma_r}(|x_p - x_q|) x_q \tag{3}$$

x_p and x_q are the value at pixel positions p and q of source image, $\| p - q \|$ is the Euclidean distance between pixel locations p and q. Parameters σ_s and σ_r will specify the amount of filtering for the image x. Equation 3 is a normalized weighted average where G_{σ_s} is a spatial Gaussian weighting that diminutions the impact of distant pixels, G_{σ_r} is a range Gaussian that reduces the impacts of pixels q when their intensity values differ from x_p. Followed by this, we apply histogram equalization to improve the contrast by adjusting the intensities of the pixel as defined in Eq. 4.

$$H = Hist(g(i, j)) \tag{4}$$

For *Category 3(C3)*, image blending [13] is performed using raw face images detected using Haar and LBP classifiers and histogram equalized images to improve the quality by diminishing the intensity variations effect as defined in Eq. 5. Here, we blend the raw face image, f and histogram equalized image, H to obtain improved quality image Q as defined in Eq. 5, where C is the blending ratio that controls the effect of

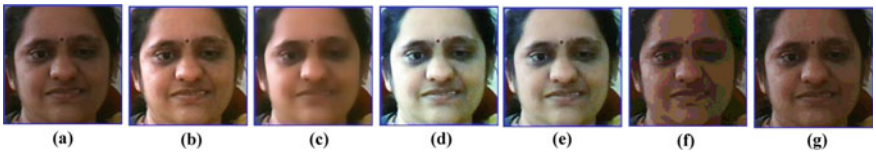


Fig. 5 Examples of image categories, **a** raw image, **b** contrast enhanced image, **c** bilateral filtered image, **d** histogram equalized image, **e** blended image, **f** quantization to 8 level, **g** quantization to 16 levels

each input image in the output. In our study, empirically we found that when setting C value to 0.3 accuracies were improved.

$$Q(I, j) = C * f(i, j) + (1 - C)H(i, j) \tag{5}$$

Later, we perform quantization by converting a range of continuous values into a finite range of discrete values, thus digitizing the amplitude values of the given image on raw images, histogram equalized images, and blended images. In *Category 4(C4)* and *Category 5(C5)*, raw images used in *C1* are considered and quantized to 8 and 16 gray levels, respectively. Similarly, in *Category 6(C6)* and *Category 7(C7)*, histogram equalized images obtained after *Category 2(C2)* are quantized to 8 and 16 gray levels, respectively. And in *Category 8(C8)* and *Category 9(C9)*, images obtained after *Category 3(C3)*, i.e., blended images, are considered and quantized to 8 and 16 gray levels, respectively. Figure 5 displays sample raw images, preprocessed images and quantized images.

3.4 Local Binary Patterns Histogram (LBPH) Recognizer

After preprocessing, the local binary patterns histogram (LBPH) algorithm is used for the recognition of faces. The general LBP operations construct the local representation of texture by comparing each pixel with the neighboring pixels in given grayscale image as shown in Fig. 4. For recognition of faces, we compare the histograms of the images using Euclidean distance (ED) as defined in Eq. 6. The evaluated distance is used as the confidence measurement. A lower confidence value indicates that the distance between the histograms is closer, and hence, the faces are matched.

$$ED = \sqrt{\sum_{i=1}^{n}(Hist1_i - Hist2_i)^2} \tag{6}$$

where $Hist1_i$ and $Hist2_i$ are the histograms of the two images. Finally, we extract LBPH features of images and separate models $(M1, M2, M3, \ldots, M9)$ are generated for each category of images.

4 Result Analysis and Discussions

To demonstrate the efficiency of our proposed face recognition model, we performed face recognition utilizing face images under varying illumination conditions which were detected using Haar and LBP classifiers. Face recognition rates (accuracy) were evaluated on various categories of images as specified in Table 1 and were also extended to resized images of size 128 × 128 to maintain uniformity in the resolution of images. The experiments were conducted using Intel(R) Core(TM) i5-8400 processor with 16GB RAM.

Figures 6, 7 and 8 show the face recognition rates obtained on faces detected using Haar and LBP classifiers on different categories of images. We perform face recognition using the LBPH algorithm on raw images, histogram equalized images, and blended images.

When face recognition was performed on raw images of varying resolutions and resized images, we observe that faces detected using LBP classifiers have outperformed with an accuracy of 98.33% and 100%, respectively, when compared to Haar classifiers. Similarly, when the quality of the face images are reduced by applying quantization of 16 levels, the proposed method achieved accuracy of 98.5% and 99.19%, respectively.

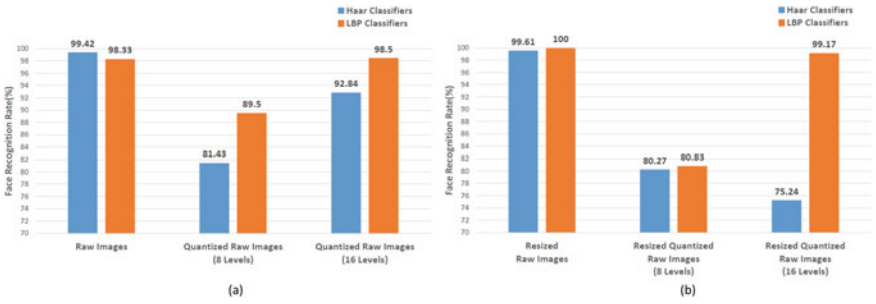


Fig. 6 Face recognition rates (%) for faces detected using Haar and LBP classifiers, **a** raw images, **b** resized raw images

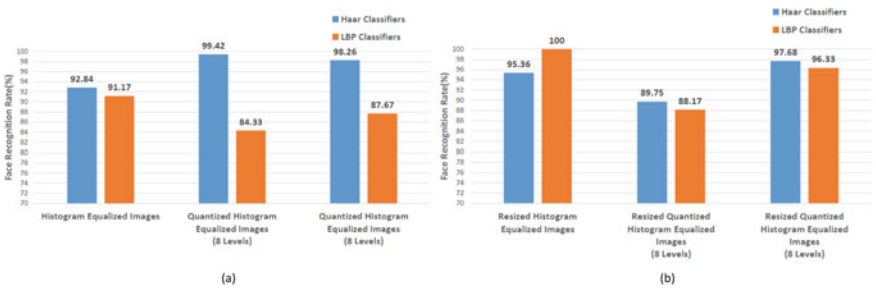


Fig. 7 Face recognition rates (%) for faces detected using Haar and LBP classifiers, **a** histogram equalized images, **b** resized histogram equalized images

Fig. 8 Face recognition rates (%) for faces detected using Haar and LBP classifiers, **a** blended images, **b** resized blended images

However, when image preprocessing techniques such as contrast adjustment, bilateral filtering, and histogram equalization were applied on the raw face images, we observed that faces detected using Haar classifiers performed better.

It was observed that when resized raw images and histogram equalized face images were blended, face recognition rate of 100% was obtained. Also, the results show that even though the quality of the resized blended images were reduced by applying quantization of 16 levels the proposed method was able to achieve an equal performance.

Based on these results, we can say that by resizing and preprocessing the images, the model was able to extract the facial features accurately. Also, it is evident from the results that even though low-quality images store less information, 100% face recognition rate has been achieved. Face detected using Haar classifiers and LBP classifiers contributed equally in recognition of faces on quantized blended image category. Thus, applying quantization of 16 levels on faces detected using LBP classifiers, we can recognize faces efficiently.

To verify the robustness of the proposed method, we have performed face recognition under varying illumination conditions using the model generated using blended

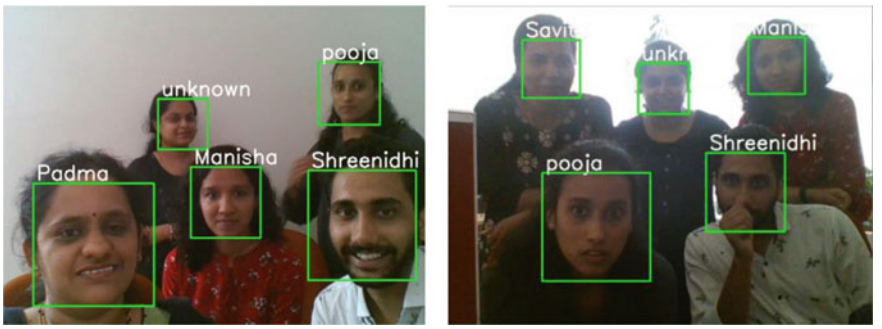


Fig. 9 Face detected and recognized under varying illumination conditions

images which were resized to 128 × 128. Sample results obtained at two different illumination conditions are depicted in Fig. 9. In spite of varying illumination conditions, the proposed method was able to recognize faces accurately. The subjects that were not involved during training were recognized as 'unknown' as shown in Fig. 9.

5 Conclusion

In this work, face detection has been performed using Haar and LBP classifiers, and recognition of those faces was performed using an LBPH recognizer. Also, promising results were obtained on the preprocessed images quantized at 16 levels. This suggests that the proposed method can be used effectively to recognize faces in low-quality images that store less information as compared to raw images. Alongside, it was observed that our proposed method using LBPH recognizer was able to recognize faces accurately under varying illumination conditions. As part of future work, face recognition can be applied to face images with different occlusions, which is one of the challenging tasks.

References

1. Yang M-H, Kriegman DJ, Ahuja N (2002) Detecting faces in images: a survey In: IEEE Trans Pattern Anal Mach Intell 24:34–58
2. Chellappa R, Manjunath BS, Malsburg CVD (1992) A feature based approach to face recognition. In: IEEE conference on computer vision and pattern recognition, pp 373–378
3. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24:971–987
4. Kasinski A, Schmidt A (2010) The architecture and performance of the face and eyes detection system based on the Haar cascade classifiers. Pattern Anal Appl 13:197–211
5. Chintalapati S, Raghunadh MV (2013) Automated attendance management system based on face recognition algorithms. In: IEEE international conference on computational intelligence and computing research. IEEE, pp 1–5
6. Liu L, Fieguth P, Guo Y, Wang X, Pietik"ainen M (2017) Local binary features for texture classification: taxonomy and experimental study. Pattern Recognit 62:135–160
7. Sharma R, Patterh MS (2015) A new hybrid approach using PCA for pose invariant face recognition. Wireless Personal Commun 85:1561–1571
8. Surekha B, Nazare KJ, Raju SV, Dey N (2017) Attendance recording system using partial face recognition algorithm. In: Intelligent techniques in signal processing for multimedia security. Springer, pp 293–319
9. Mady H, Hilles SMS (2018) Face recognition and detection using Random forest and combination of LBP and HOG features. In: International conference on smart computing and electronic enterprise (ICSCEE). IEEE, pp 1–7
10. Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: IEEE computer society conference on computer vision and pattern recognition, vol 1. IEEE
11. Szeliski R (2010) Computer vision: algorithms and applications. Springer Science & Business Media

12. Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: International conference on computer vision. IEEE, pp 839–846
13. Vision EDM et al (1990) Theory, algorithms and practicalities. Davies (Author)

A Bibliometric Analysis of Green Computing

Arti Chandani, Smita Wagholikar, and Om Prakash

Abstract The objective of this paper is to study the trends in the research papers in last two decades in the area of green computing. The present study is aimed to find the existing literature in the field of green computing by undertaking a bibliometric review. This work examines the research papers which were extracted from Scopus database. Scopus is one of the largest databases, and it is much user-friendly in terms of extracting the existing studies undertaken by other researchers. The papers extracted from Scopus were from year 2001 to 2020, and authors have used only one filter in addition to year, i.e. language as English. The authors have used VOSviewer software to generate network diagrams. The results show that there has been a significant increase in the number of research studies since 2012 onwards. Around 84% of research studies are published in the conference proceedings and journal. Another important finding was that India was at the top of the list in terms of number of articles by a country where India has 186 research papers in the area of green computing. The study also gives an insight into the co-authorship and keyword analysis. The study is very important from the perspective that it gives a direction to the future researchers by apprising the existing studies in the area of green computing. The study is unique in a way that it gives co-authorship and keyword analysis using VOSviewer.

Keywords Green computing · Bibliometric analysis · Scopus · Co-authorship · Keyword analysis

A. Chandani (✉)
SIMS, Symbiosis International University, Pune, Maharashtra 411028, India
e-mail: arti.chandani@sims.edu

S. Wagholikar
SIT, Symbiosis International University, Pune, Maharashtra 411028, India

O. Prakash
SICSR, Symbiosis International University, Pune, Maharashtra 411028, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_52

1 Introduction

Computing with a focus on environmental concerns and energy efficiency has given rise to a new area for research and application known as green computing. The release of carbon dioxide and energy is ever increasing while computers and computing are being used in every aspect of our lives. Green computing is also referred to as green IT. The designing and operating of computers in a way that is eco-friendly and sustainable is the requirement of the current situation. Another aspect of this is to recycle and reuse computer assets and spread awareness about the need for green computing [2]. The awareness to become green should be a part of the curriculum for students [16]. This will ensure that upcoming programmers will use future computing technology keeping the eco-friendly angle in mind [8] and also gain cognizance of the management of e-waste through EPR [7].

The use of cloud computing shows an increasing trend which demands a larger infrastructure for the data centre. This in turn leads to a generation of greater carbon emission, use of energy and increases the expenses. In this regard, techniques for saving energy saving in servers becomes a requirement. To maintain performances of cloud computing and optimise the energy consumption, [4] reported the use of green cloud alternatives for the management of resources in data centres. This was beneficial in reducing the requirement for energy and is a step towards achieving green computing.

The importance and awareness of going green in computers was cited in 2001 when there was very little research reported. As global warming became a matter of serious concern because of the unplanned use of resources including infrastructure and machines, the necessity to introduce green computing, cloud computing and smart networking gained prime importance. There was a rapid increase in the realisation of recycling, reuse, keeping a track of low carbon emissions and e-waste generation [18] and sustainability in order to address these concerns. Research in green computing showed a continuous increasing trend from the last ten years. There is a steady increase in the publication count from 28 in 2010 to 88 in 2020 as seen from Scopus database search.

With the advent of green IT and computing technology and its ever-increasing need, various methodologies, modified operations, algorithms and experiments were carried out to decrease the consumption of energy and emission of greenhouse gases. The present bibliometric study refers to a systematic study of the growth of awareness and application of computing the environmental friendly way. This bibliometric study records the milestones in developing and implementing computing techniques that move towards safe, clean and green computing. The literature from Scopus database search includes reviews, overview, survey and book chapters; however, a bibliometric analysis is not available.

This domain of computing where computing could be done by using minimum energy with the same performance or outcome is a simple way to explain green computing. Green computing has its major application in computer science (CS),

and this can be seen from the articles published. The maximum number of articles are published in CS numbering to total of 512.

2 Literature Review

In a recent report over cloud computing, an ant colony system algorithm was applied to reduce energy consumption over idle virtual machines. Hossain et al. [9] reported that the algorithm not only helps in energy efficiency but also deals with another problem of virtual machine placement. This considerably reduces costs over resources and energy in a data centre. While computers are being made, a lot of water, chemicals and fossil fuels are utilised. Also, a considerable amount of carbon dioxide is released when computers are being used. Arya et al. [1] discussed that keeping the computer screen power off when unused reduces the energy consumption. Besides, pollution and energy can be significantly reduced when file down loading is migrated to server from the host machine. Naren and Dishita [15] mentioned a technique that is applied to measure the actual power drawn from the mains and the power utilised so that a difference between the two can be calculated to reduce wastage of power. The research reports the use of an analytical instrument that measures efficacy of a CPU and its performance to minimise energy wastage.

Green computing aims at making computing eco-friendly by using more biodegradable products, minimum energy and reducing waste. Jacob and Preetha [10] in their research discussed going green in computing by the application of conventional computing methods along with reducing the requirement of power. The use of a smartphone as a support device is mentioned in the study. Germain et al. [5] reported the unique facility of Green Computing Observatory for green IT processes. This facility improves energy efficiency by compiling data on energy and its analysis to move a step forward towards green computing.

Gholipour et al. [6] reported the introduction of a modern service delivery named container for managing IT services and resource management. Besides being portable and scalable, containerization enables considerable reduction in use of energy and violation of service-level agreement leading to a new development in green computing. These can be achieved by implementing use of a joint virtual machine and container migration by using a simulator known as a ContainerCloudism. In a recent research, Mandal et al. [13] reported that excessive migration leads to increase in energy and operational cost which in turn lead to increase in carbon footprint. To help the decrease in use of power a policy that assists selection of a virtual machine for migration is suggested. This technology is an important contribution in making energy-aware cloud computing.

Bharani and Jayalakshmi [2] reported methodologies in cloud computing which are oriented to reduce environmental hazards. Various methods such as containerization, green network and protocols, and designing eco-friendly virtual machines were included in the research. The researchers also mentioned that implementing the policy of reuse, reduce and recycle is an effective method to attain software development

which is both environment-friendly and sustainable. While multi-access computing deals with computing tasks, it is also known to increase the consumption of volatile energy. In order to reduce the high energy demand and supply, Munir et al. [14] developed a three-step process that increased the performance and renders a policy that is best for energy profile.

Lin and Wang [12] proposed a new model called quality of green to avoid high energy consumption of the network. Fuzzy mathematics is used as a base to make a comprehensive evaluation matrix. Liu reported the use of network performance along with traffic and topology to understand the connection between quality of service and the index for green energy consumption. Davis and Raianu [3] brought to our notice that there is a nexus between mathematics and elementary computer graphics and mentioned how a software planimeter can be prepared. The application of various integration methods in this research may be an example for students studying mathematics making the study very interesting.

Liao et al. [11] reported development in the server to reduce the cost for decryption to user's and energy requirements for the benefit of the environment. In a recent research to enhance energy efficacy, Scionti et al. [17] mentioned the use of a novel resource scheduler. The research also mentioned the implementation of an algorithm to improve scheduling.

2.1 Research Objectives

1. To study the overall scenario of green computing and recent developments in computing as an effort towards eco-friendly
2. To find out top ten cited authors, organisations and countries
3. To identify the possible areas and domains of green computing that can be explored and suggestions for futuristic developments.

3 Research Design

The present research is aimed to understand the existing knowledge in the area of green computing using bibliometric analysis. We have used the Scopus database for extracting the articles indexed in Scopus. The search word used was green computing, in article title, which resulted in the articles 793. We applied filters to extract the articles and one of the filters was the time period. The articles for the period of 2001–2020 were included in the research. The language "English" used another filter. Using these filters, total articles were 719 which were downloaded from Scopus in csv format. Once the file was downloaded, we checked the articles, and there were some missing entries and duplicate entries which were cleaned. So the final number of articles were 708. VOSViewer was the tool which was used by us in order to generate network diagrams, and the csv file which was taken from Scopus was used

as the source data for this. The results of the network diagram are presented in Sect. 3 under the data analysis section.

4 Data Analysis

The present section deals with the analysis of the articles which were extracted from Scopus.

The number of articles published are shown in Table 1 for the period of 20 years starting from 2001 to 2020. It can be seen that there is a surge in the number of articles published from the year 2010 onwards. Maximum number of articles published were in the year 2020 which were 88, and this shows the increasing importance of the subject under study (Table 2).

It is interesting to note that the number of articles published in conference proceedings and journals was almost the same as 299 and 298, respectively, and the number of papers published in book series and books is very less.

The top two institutes who are having maximum number of articles are from India where Vellore Institute of Technology is having 11 articles followed by Amity University having 10 articles. Table 3 shows the top ten affiliations and institutions having articles in the area of "green computing" (Table 3).

Table 1 Publication count between years 2001 and 2020

Year	Count	Year	Count	Year	Count
2001	4	2008	8	2015	54
2002	1	2009	10	2016	69
2003	2	2010	28	2017	53
2004	3	2011	38	2018	86
2005	2	2012	50	2019	78
2006	3	2013	64	2020	88
2007	5	2014	62		
				Total	660

Table 2 Source type and Count

Source type	Number of documents
Conference proceedings	299
Journal	298
Book series	69
Book	36
Trade journal	6
Total	708

Table 3 Top ten Affiliations and Count

Affiliation	Number of documents
Vellore Institute of Technology	11
Amity University	10
Zhejiang University	8
Universiti Teknologi Malaysia	8
University of Pittsburgh	7
Kyung Hee University	6
Arizona State University	6
University of Houston	6
CNRS Centre National de la Recherche Scientifique	6
Universidad de Chile	6

Table 4 Top ten countries and count

Affiliation Country	Number of documents
India	186
The United States of America	130
China	112
The United Kingdom	37
France	31
Canada	25
Australia	22
Malaysia	22
Italy	21
Saudi Arabia	18

Another analysis was done to identify the countries which have publication in the area of "Green Computing". It is interesting to know that the highest number of publications are from India where the total number of articles are 186 followed by the USA with 130 articles as given in Table 4.

We have used VOSviewer to draw the network diagrams for the purpose of bibliometric analysis. Figure 1 shows the co-authorship of authors. There were 1785 authors of the data extracted from Scopus using the filters as given in the research design section. We have used 2 documents as the minimum number of documents and 1 citation as the minimum citation of an author. The network diagram shows that Wu. J has the strongest link among the selection with 40 as the total link strength and 14 documents which was followed by many authors who had total link strength of 30. Wu. J. also had the maximum number of documents as 14 followed by 10 which were authored by Zhang. H with a total link strength of 30.

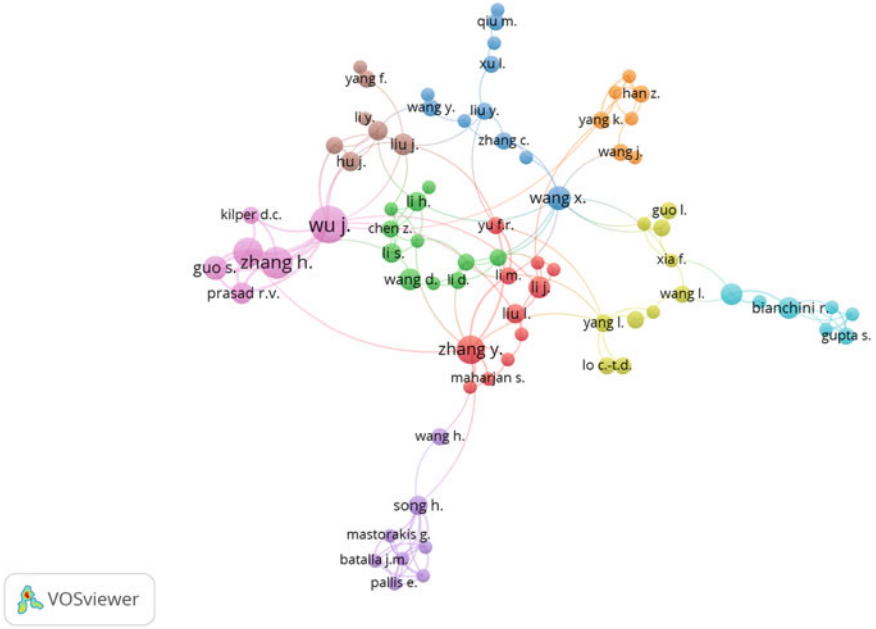


Fig. 1 Co-authorship of authors

Figure 2 shows a network diagram of authors where the unit of analysis is country. The sample data has authors from 78 countries, and we have used minimum 1 document and minimum 1 citation, and the resultant diagram is shown in Fig. 2. The maximum link strength is China with a total link strength of 82 followed by 77 of the USA. If we see the maximum number of documents with respect to country, India has 186 documents followed by USA with 130 documents and China with 112 documents; while in terms of citations, maximum citation are 1734 for the USA, followed by 1711 of China and 983 of India. This shows that in terms of green computing, the majority of the research is being carried out in China, the USA and India (Figs. 2 and 3).

Figure 3 shows the total number of keywords were 4278 for the sample data while we have used the 2 as the minimum number of occurrence of keywords which results in 975 keywords. The highest link strength was of the word "Green Computing" with 3484 followed by "Cloud Computing" as 2218. In terms of occurrence, also, "Green Computing" and "Cloud Computing" were high with a number of 378 and 208 occurrences, respectively. This also highlights that when researchers are working in the area of green computing, the words green computing and cloud computing are used much. The other keywords which are also being used by researchers are energy efficiency and energy utilisation among others.

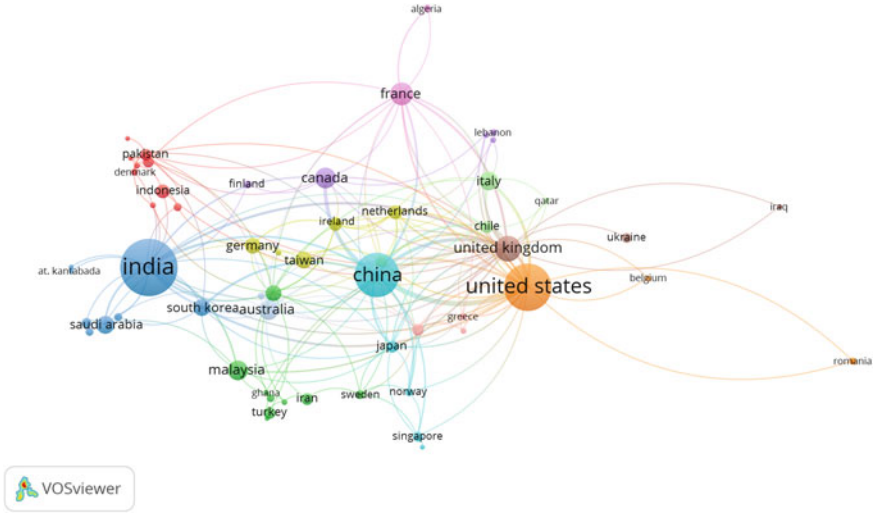


Fig. 2 Co-authorship of authors using country

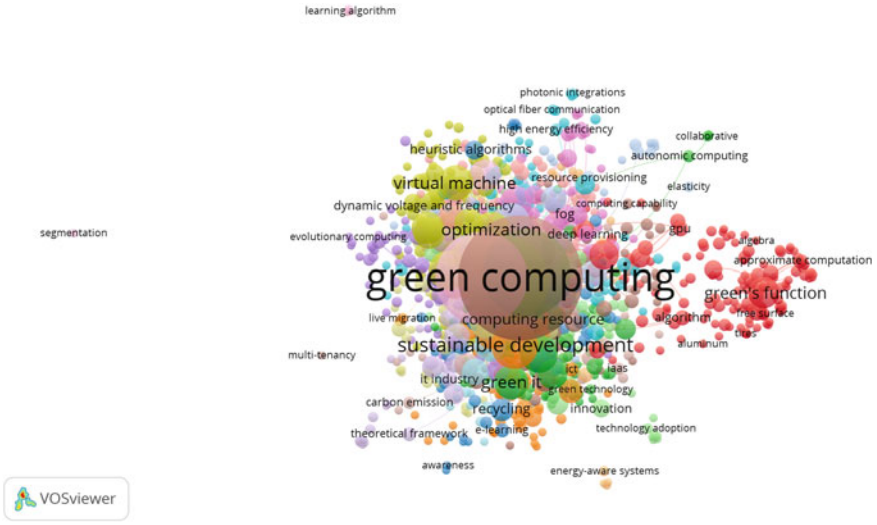


Fig. 3 Co-occurrence of all keywords

5 Conclusion

We have performed a bibliometric analysis, using VOSviewer software, of the research in the area of green computing. The study covered the research articles, paper published and indexed in Scopus database in the field of green computing, from 2001 to 2020. The number of studies were in single digit till 2008, post which there was gradual increase in the paper while number of paper increase significantly from 2012 onwards. The studies clearly show that there is an increasing interest in the topic and more so due to increase in the environment consciousness among all. The study and analysis shows that India and the USA are the leading countries in terms of publication of research paper with 186 and 130 papers, respectively. China is at number three with 112 papers in the area of green computing. Another interesting finding of the study is that most number of documents by an institution are from 2 Indian institutions where Vellore Institute of Technology has 11 documents and Amity University has 10 documents. This was followed by Zhejiang University, China and Universiti Teknologi, Malaysia where both of these universities had 8 documents each.

5.1 Limitations

The present study is limited by the various filters used in the study which were practical, however, can be considered as the limitation of the study. The first and foremost limitation of the study emanates from the fact that "Green Computing" was the key word which was used to search and extract the articles from Scopus. If there were more keywords, the results would have been different; however, due to the objective of the study, it was limited to "Green Computing" only. Another limitation comes from database, i.e. Scopus. We have used the Scopus database which is one of the largest databases of articles; however, it has its own limitation. We have not considered other databases such as Web of Science or others in order to limit the scope of work. We have used VOSViewer as the tool for the network analysis, which could be another limitation. Another limitation of the study is that the recent studies might not have been included in the study as author used the filter of year 2020 while that was the objective of the study as well.

5.2 Implications

The study is very useful from the perspective of research as well as from a managerial perspective. This study gives a landscape of the research in the area of green computing and also gives the areas in which the research has not been done which can be taken up by the future researchers working in the field of green computing.

5.3 Scope for Future Work

As the green computing trend continues to gain impetus, there is scope for exploring the use of biodegradable computer equipment and creating more awareness about reducing the power utilisation by switching off machines when idle and making algorithms that will make use of optimum energy while keeping the performance in its place. There is also a greater opportunity to educate the coming generation and budding engineers by introducing their curriculum with eco-friendly aspects of green IT. This will certainly pave the way for technology with biodegradable and energy efficiency. There can be more policies implemented for making green computing mandatory to conserve our planet earth.

References

1. Arya MS, Manjare A, Naik P, Sheikh KH (2020) A green computing technique for reducing energy consumption for desktops using proxy downloading. In: Information and communication technology for sustainable development. Springer, Singapore, pp 823–831
2. Bharani Dharan G, Jayalakshmi S (2020) Harnessing green cloud computing—a energy efficient methodology for business agility and environment sustainability. Int J Emerg Trends Eng Res 8(8):4193–4200
3. Davis P, Raianu Ş (2007) Computing areas using Green's theorem and a software planimeter. Teach Math Appl: Int J IMA 26(2):103–108
4. Diouani S, Medromi H (2018) Green cloud computing: efficient energy-aware and dynamic resources management in data centers. Int J Adv Comput Sci Appl 9(7):124–127
5. Germain-Renaud C, Furst F, Jouvin M, Kassel G, Nauroy J, Philippon G (2011) The Green computing observatory: a data curation approach for green IT. In: 2011 IEEE ninth international conference on dependable, autonomic and secure computing. IEEE, pp 798–799
6. Gholipour N, Arianyan E, Buyya R (2020) A novel energy-aware resource management technique using joint VM and container consolidation approach for green computing in cloud data centers. Simul Model Pract Theory 104:102127
7. Hanief S, Kartika LGS, Srinadi NLP (2017) A survey regarding the readiness of campus in Indonesia on the adoption of green computing. In: 2017 Second international conference on informatics and computing (ICIC). IEEE, pp 1–6
8. Hernandez AA (2019) An empirical investigation on the awareness and practices of higher education students in green information technology: Implications for sustainable computing practice, education, and policy. Int J Soc Ecol Sustain Dev (IJSESD) 10(2):1–13
9. Hossain MK, Rahman M, Hossain A, Rahman SY, Islam MM (2020) Active & idle virtual machine migration algorithm—a new ant colony optimization approach to consolidate virtual machines and ensure green cloud computing. In: 2020 Emerging technology in computing, communication and electronics (ETCCE). IEEE, pp 1–6
10. Jacob JS, Preetha KG (2012) A novel approach for green computing through event-driven power aware pervasive computing. In: International conference on computer science and information technology. Springer, Berlin, pp 81–88
11. Liao Y, Zhang G, Chen H (2020) Cost-efficient outsourced decryption of attribute-based encryption schemes for both users and cloud server in green cloud computing. IEEE Access 8:20862–20869
12. Lin H, Wang J (2012) A new evaluation model on green low carbon energy network for cloud computing—quality of green (QoG). In: 2012 IEEE Asia Pacific cloud computing congress (APCloudCC). IEEE, pp 1–6

13. Mandal R, Mondal MK, Banerjee S, Biswas U (2020) An approach toward design and development of an energy-aware VM selection policy with improved SLA violation in the domain of green cloud computing. J Supercomput 1–20
14. Munir MS, Abedin SF, Tran NH, Han Z, Hong CS (2019) A multi-agent system toward the green edge computing with microgrid. In: 2019 IEEE global communications conference (GLOBECOM). IEEE, pp 1–7
15. Naren T, Dishita B (2017) TVAKSHAS-an energy consumption and utilization measuring system for green computing environment. In: International conference on ubiquitous communications and network computing. Springer, Cham, pp 37–45
16. Qian K, Lo CTD (2010) Leveraging CS capstone project and green smart energy computing with WSN in a box. In: Proceedings of the 48th annual southeast regional conference, pp 1–2
17. Scionti A, D'Amico C, Ciccia S, Li Y, Terzo O (2019) Smart scheduling strategy for lightweight virtualized resources towards green computing. In: Conference on complex, intelligent, and software intensive systems. Springer, Cham, pp 305–316
18. Sharma H, Singh JB (2020) The effectual real-time processing using green cloud computing: a brief review. In: 2020 9th international conference system modeling and advancement in research trends (SMART). IEEE, pp 275–280

Fractal Image Coding-Based Image Compression Using Multithreaded Parallelization

Ranjita Asati, M. M. Raghuwanshi, and Kavita R. Singh

Abstract Fractal image coding-based image compression is characterized by its high compression ratio, high-resolution, and lower decompression time. In spite of these advantages, it is not being widely adopted because of its high computation time. Attempts made to reduce the computation duration in fractal image compression (FIC) fall into two categories like heuristics-based search time reduction and parallelism-based reduction. In this work, we have proposed a multithreading-based parallelism technique on the multi-core processors to minimize the compression duration. The compression duration of the suggested multithreading process is tested upon the images having different resolutions. It is observed that the proposed solution has reduced the compression time by almost 2.51 times as compared to sequential method.

Keywords Fractal image compression · Encoding · Fractal image coding · Data parallelism · Multithreading

1 Introduction

Image compression is the technique of encoding an image to lessen the image size which results into efficient storage and transmission. Compression schemes exploit the irrelevant and redundant information in the images [1]. Existing compression techniques are of two types: lossless and lossy. There is no information loss and the compressed image retains its quality in case of lossless methods. There is a bearable information loss in lossy compression methods to achieve higher compression ratio. Fractal image coding is a lossy compression scheme and aimed on reducing the redundancy in the self-similar images [2]. It achieves higher compression ratio

R. Asati (✉) · K. R. Singh
Department of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, India
e-mail: ranjita.asati@gmail.com

M. M. Raghuwanshi
Department of Data Science, G.H. Raisoni College of Engineering and Management, Pune, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_53

by losing some information in the compressed image. FIC permits an improbable volume of data to be kept in greatly encoded records. A fundamental feature of FIC is to convert images into fractal codes. This is identified as fractal scaling, ceasing original pixel structure. An image is converted forming a group of contractive similar transformations exploiting the self-affinity of the image. The parameters of these transformations are stored as compressed file [3]. FIC is able to provide higher compression ratio [4]. Higher the compression ratio, higher the computation complexity in the fractal coding schemes due to search of matching blocks [5].

Numerous efforts are made to decrease the intricacy of the fractal image compression (FIC). The attempts are divided into two categories; in first category, efforts are made to reduce the search time for matching blocks (algorithm-based approach), and the second category is focused on the use of modern computer architecture for speed up (hardware-based approach). The proposed work falls into second category that uses multithreading parallelism on multicore processors to decrease the compression time of an image. In this work, we make use multithreading-based parallelization on multicore processors to diminish the encoding duration. The search time for localization of approximate range block for each domain block can be reduced, and localization process can be done with least error due to this multithreading parallelization. Remainder of this paper is structured as follows: Sect. 2 put forward literature review on the fractal image coding techniques. Section 3 presents the suggested approach with its algorithm. Section 4 discourses the result part. In the end, some conclusions are presented in Sect. 5.

2 Related Work

FIC is a lossy compression technique developed by Barnsley [BH86] and Jacquin [2] to compress an image by encoding it as a transformation. Firstly, it divides the original image into domain blocks (they can be any size or shape). Then, a collection of possible range blocks is defined. For each domain block, the algorithm then examines for a suitable range region that, when applied with an proper affine transformation, very closely look like the domain block. Later, a Fractal Image Format (FIF) file is generated for the image. This file contains information on the choice of domain regions, and the list of affine coefficients (i.e., the entries of the transformation matrix) of all allied affine transformations. This process is very time-consuming, especially during the search for the proper range regions. But once the compression is done, the FIF file can be decompressed very rapidly. The literature survey on existing works is studied in two categories of heuristics-based search time reduction and parallelism-based reduction.

2.1 Heuristics-Based Search Time Reduction

Ismail et al. [6] proposed a hybrid technique using variable range block and optimum domain block to decrease the compression duration in FIC. The solution is based on hexagonal structure of spiral architecture. The reduction in encoding time is achieved by reducing the number of matching operations and substituting its equivalent results with lower mean square error. Borkar et al. [7] reduce the encoding time in FIC using wavelet transform. In addition to reduction in time, use of wavelet also helps to enhance the visual appearance. Nadia et al. (2017) suggested a novel approach using fractal dimension to reduce the search process for each range block. This approach makes a small compromise in image quality to decrease the compression time. Wang et al. [8] presented a practice to accelerate FIC using Pearson's correlation coefficient. Sparse searching is implemented to find the matching domain region; due to this, encoding time is reduced at a cost of minor compromise in the quality of image. Hsu et al. [9] proposed a repetition-free FIC. A domain pool constitutes the mean image for both encoding and decoding process. This domain pool is able to avoid the iteration process. The method has lower encoding and decoding time, but it compromises the PSNR. Cao et al. [10] proposed a novel orthogonal sparse FIC to improve compression time. This method takes into account image texture feature. Most of these works achieved reduction in compression time by compromising the PSNR.

2.2 Parallelism-Based Reduction

Parallel computing is a computing where the tasks are fragmented into distinct parts that can be accomplished simultaneously. Every part is again fragmented into a series of instructions. These instructions are then carried out simultaneously on different processors. Parallel systems make simultaneous use of multiple computer resources that can employ a single computer with multiple processors, a computer network forming a parallel processing cluster or a combination of both. The heart of parallel processing are CPUs. Flynn's taxonomy is a peculiar arrangement of computer architectures dependent on the number of parallel instruction (single or multiple) and data streams (single or multiple) available in the architecture.

Min et al. [11] achieved speed up of FIC using a massively parallel implementation on a pyramid machine. Authors could diminish the computation intricacy from $O(n^4)$ to $O(n^2)$ for an $n \times n$ image using the parallel architecture. Erra et al. [12] proposed a parallel FIC using programmable graphics hardware. The solution used SIMD architecture to speed up the base line approach of fractal encoding. Palazzari et al. [13] exploited the massively parallel processing on SIMD machines to solve the problem of higher encoding time in FIC. Lee et al. [14] suggested a parallel quad tree-based FIC. It performs the fractal image coding based on quad tree partitioning. Hufnagl et al. [15] proposed a parallel version of FIC over massively parallel SIMD

arrays. The speed up is achieved in encoding time based on load balancing over 2D SIMD arrays. Bodo et al. [16] proposed a quad tree-based FIC scheme. The encoding time is reduced in this parallel architecture by maximizing the processor utilization. Haque et al. [17] carried out sequential as well as parallel FIC in Compute Unified Device Architecture (CUDA) platform. Medical images were used for experimentation achieving 3.5 times more speed up compared to CPU. Saad et al. (2016) proposed a parallel architecture using FPGA for implementing a full search FIC. A near optimal performance is achieved using low-cost FPGA hardware. Malik et al. [18] proposed FIC algorithm employing deep data pipelining for high-resolution images. By limiting the search in neighboring blocks, reduced further encoding time. The solution can compress large size images in less time.

Though these parallelism-based solution are able to achieve significant speed up, but they require expensive hardware's.

3 Proposed Solution

The proposed solution for achieving speed up in FIC based on the concept of adding multithreading feature to the original FIC encoding and decoding process so that it becomes more suitable for the data level parallelism without much inter thread communication. The serial version of FIC is modified to incorporate multithread parallelism. In the serial FIC, the initial image of dimension $M * M$ is split to m non-overlapping range blocks of size $r * r$ where $m = M/r^2$ and n overlapping domain blocks of size $2r * 2r$ where $n = (M - 2r + 1)^2$.

Fractal coding follows Iterated Function System. In first step image is partitioned into non-overlapping blocks called range blocks and overlapping sub-blocks called domain blocks. Each range block is mapped with one of the available domain blocks. To locate most suitable matching domain block, a range block has to be compared with each domain block to record the minimum distortion. Eight affine transformations do not change pixel values; they simply shuffle pixels within a range block, in a deterministic way. For every range block, search is done on the domain pool for a block which can be mapped linearly to the range block with smallest error. As a result, an approximate domain block and a suitable contractive affine transformation is selected according to the minimization function below

$$d(R_i, w_{ik}(D_k)) = \min d\left(R_i, w_{ij}\left(D_j\right)\right) \tag{1}$$

where w_{ik} is the contractive affine transformation from D_k to R_i. This is done in such a way that mean square error distance (represented as $d(R_i, w_{ij}\left(D_j\right)$ from range block R_i and the transformed domain block $w_{ij}(D_j)))$ is minimized.

Each thread processes a range block in the range block pool. The thread does transformation on domain blocks, finds the most suitable transformed blocks based on least mean square error between the transformed domain block and range block.

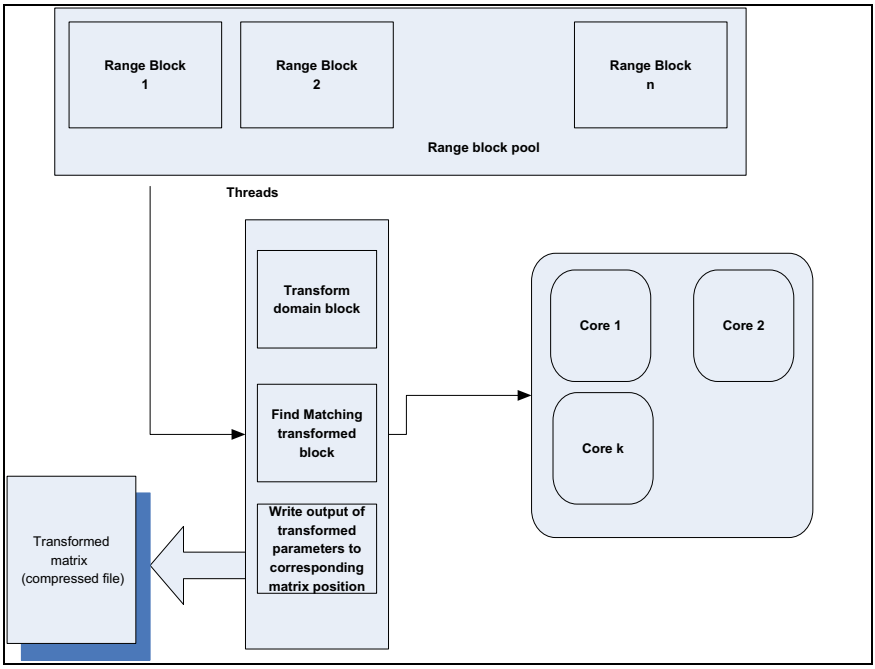


Fig. 1 Proposed architecture

The threads are scheduled for execution on the any cores of the multi core processor by the OS. Since there is no inter thread communication, the usual pitfalls of multithread communication like resource contention, racing, deadlocks etc. are totally avoided. Each thread generates the transformed parameters for it range block which is written to the matrix and returned as a compressed file. In the proposed work, multiple threads are assigned to lookup for domain block for each range block. Each read takes the range block as input. For the range block it iterates through the domain block pool and select the domain block which on transformation can be approximated to a range block with minimum distortion. Each searching for domain block is done in parallel, it reduces the encoding time. Each thread will execute in parallel on different cores. Since processing steps for each range block use same instructions but different data, a data level parallelism can be achieved without much inter thread cooperation. Only when returning the final compressed file in terms of transformation parameters, a wait must be done to collect the results from all threads. The parallelism architecture of the proposed solution is given in Fig. 1. Both the encoding and decoding process are parallelized to achieve speed up. The steps in the parallelized encoding process are given below.

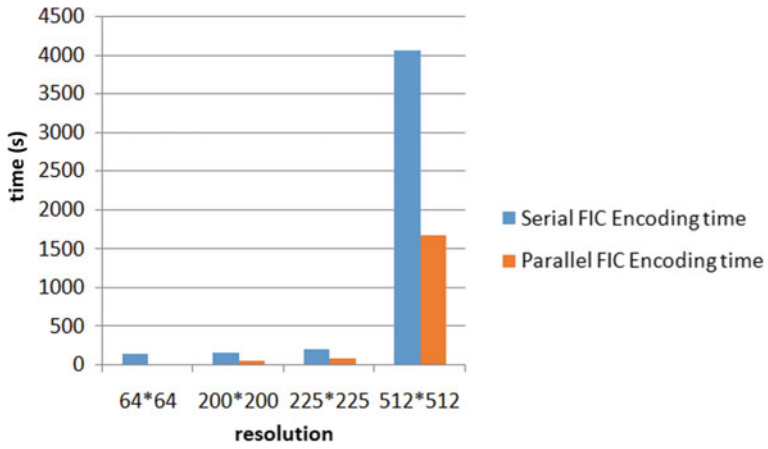


Fig. 2 Serial and parallel FIC encoding time

Algorithm: Parallel FIC Encoding

Step 1. Input image is divided into non-overlapping range regions and overlapping domain regions.

Step 2. For every range region start a thread.

In each thread for the domain region, find the approximate range region by selecting the range block with lowest MSE distance of range region and transformed domain regions.

Step 3. Wait for all thread to complete.

Step 4. Return all the transformations.

The decoding process is parallelized by splitting the compressed file into portions and doing decoding process on each portion. Due to parallel reconstruction, the time of reconstruction is reduced. The steps in decoding algorithm are given below.

Algorithm: Parallel FIC Decoding

Step 1. Generate range blocks from the encoding result file.

Step 2. Chose a starting image, which should be of same dimension as that of the initial image.

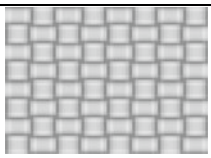
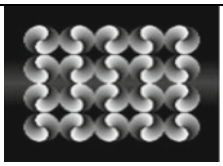
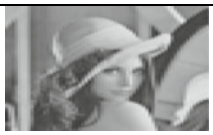
Step 3. Split range block matrix to K partition and start K thread.

Step 4. In each thread do following.

Apply the transformation parameters derived from the transformed block for each range block.

Substitute the range block pixels with the pixels which are obtained from the transformed block.

Table 1 Image used for experimentation

Image		
		
		
		
		

Both transformations and mappings are enforced on the initial image repetitively until the image is restored.

Step 5. Wait for all thread to complete.

Step 6. Return the reconstructed image.

Though only concern in FIC is encoding time, in this work, decoding process is also parallelized just to increase the resource utilization of multiple cores (Table 1).

4 Results

The performance of the proposed solution was conducted in 10th generation Intel core i5 10300H, 2.5 GHZ, 4 cores, 8 threads, 8 GB RAM. Performance was tested with 40 images in different resolutions of 64 * 64, 200 * 200, 225 * 225, and 512

Table 2 Comparison of encoding and decoding time

Size	Serial FIC	Serial FIC	Parallel FIC	Parallel FIC
	Encoding time (s)	Decoding time (s)	Encoding time (s)	Decoding time (s)
64 * 64	36.49	0.94	7.59	0.94
200 * 200	155.3	3.24	29.63	1.88
225 * 225	228.3	4.36	93.40	2.77
512 * 512	4058.1	18.84	1824.4	11.00

* 512. The dataset image for performance testing is obtained by taking 10 different images in resolution of 512 * 512 and rescaling them to dimensions of 64 * 64, 200 * 200, and 225 * 255. By this way, dataset of 40 images was created with different scales. By this way of testing the FIC against these recalled images, the robustness of the proposed FIC algorithm against different scales can be tested. These images were then used for FIC testing. Since there were no standard dataset for testing fractal image compression, we used the same set of images used for testing block complexity-based FIC proposed in [19].

The proposed solution was compared with serial version of FIC implementation with both encoding and decoding process realized as a single process. The encoding and decoding time are measured for different images, and the result is summarized in Table 2. The results for encoding time are as below in Fig. 2.

From this result, it is evident that as the resolution of the image increases, the encoding time increases, but the increment is lower in the proposed solution compared to Serial FIC. The average encoding time in Serial FIC is 1145 s, but the in the proposed solution, the average encoding time is 455 s. Thus, the proposed solution has an average speed up of 2.51 compared to Serial FIC. The average decoding time in Serial FIC is 7.14 s, but in the proposed solution, it is only 3.67. The proposed solution has a 1.95 times speed up in decoding time compared to Serial FIC. The speed up in encoding process in the proposed solution is due to parallelization in the steps of calculation of transformations and matching to every transformation to find best match with least error.

The reduction in the decoding time is due to parallelization in the step of replacing the pixels of range block from the transformed block of each range block (Fig. 3).

The box-whisker plot of encoding time for different images in four different scales of 64 * 64, 200 * 200, 225 * 225, and 512 * 512 is given in Fig. 4.

Though earlier many solutions have been proposed for FIC, they did not fully exploit the processing capability of the underlying hardware. Due to this, hardware was underutilized and it took long time for compression. Parallelization was the solution to increase the hardware utilization and reduce the time for compression. There were two options for parallelization, process level and thread level. In case of FIC, thread level parallelization was suited more than process level as the processing overhead of inter-process communication would be high for FIC, and it would ruin the advantages of parallelization. Due to this, thread-level parallelization was adopted.

The results proved that compression time is reduced by increasing the utilization of underlying hardware using threads.

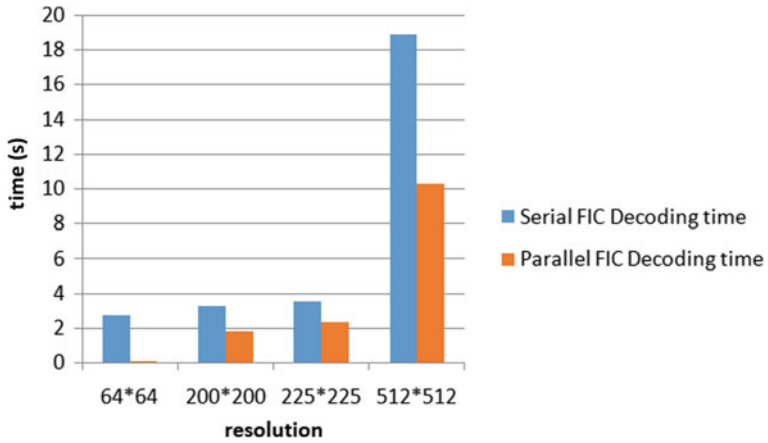


Fig. 3 Serial and parallel FIC decoding time

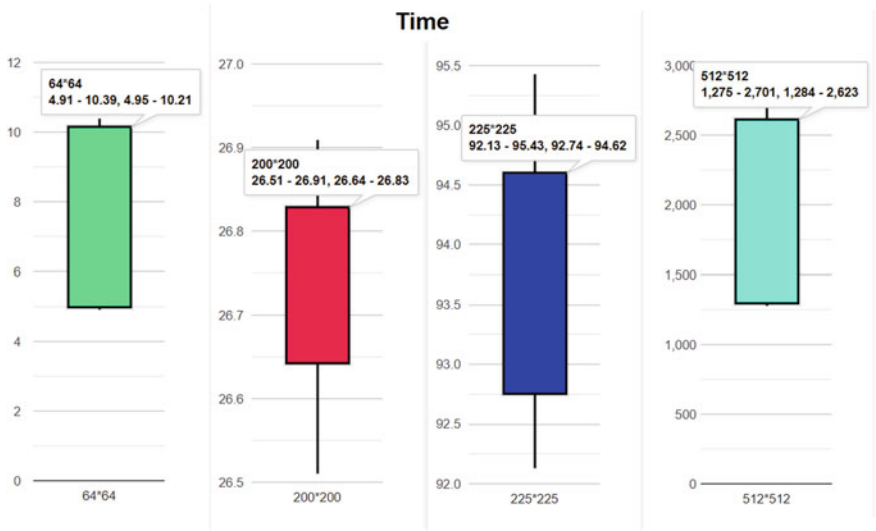


Fig. 4 Box-whisker plot of encoding time

5 Conclusions

In this work, a multi thread-based parallelization using multi-core processors is implemented to speed up the execution of FIC. FIC is generally resource and time-consuming algorithm, and the computation complexity is exponentially increasing with increase in size of the image. Though many algorithms have been proposed for FIC, they have two issues of underutilization of hardware and large time for encoding. In this work, thread-level parallelism is introduced to FIC to solve the two issues of underutilization and larger encoding time. The performance of the proposed solution was tested for different images at different scales, and results show an average 2.51-time reduction in encoding time. The reduction in encoding time was achieved at the cost of increasing the hardware utilization through thread-level parallelism. A limitation in this approach is that for self-similar blocks, the computation effort is repeated. This limitation can be solved by further parallelization by splitting the image to high-level blocks and clustering them based on entropy or structural similarity and applying parallel FIC for a representative block in each cluster and approximating other blocks in cluster based on it. This will further reduce the encoding time.

References

1. Hussain A, Al-Fayadh A, Radi N (2018) Image compression techniques: a survey in lossless and lossy algorithms
2. Jacquin A (1989) A fractal theory of iterated Markov operators with applications to digital image coding
3. Asati R, Raghuwanshi MM (2020) Fractal image compression: a review. Int J Future Gener Commun Network 13(1s):66–75
4. Wohlberg B, de Jager G (1999) A review of the fractal image coding literature. IEEE Trans Image Process 8
5. Fisher Y (1995) Fractal image compression: theory and application. Springer, New York
6. Ismail M, Reddy BTB (2016) Spiral architecture based hybrid fractal image compression. In: International conference on electrical, electronics, communication, computer and optimization techniques (ICEECCOT)
7. Borkar E, Gokhale A (2017) Wavelet based fast fractal image compression. In: International conference on innovations in information embedded and communication systems (ICIIECS)
8. Wang JJ, Chen P, Xi B et al (2017) Fast sparse fractal image compression. PLOS ONE 12(9)
9. Hsu C-C (2018) Iteration-free fractal mating coding for mutual image compression. In: International symposium on computer, consumer and control (IS3C)
10. Cao J, Zhang A, Shi L (2019) Orthogonal sparse fractal coding algorithm based on image texture feature. IET Image Process 13(11):1872–1879
11. Min X, Hanson T, Merigot A (1994) A massively parallel implementation of fractal image compression. In: IEEE international conference on image processing
12. Erra U (2005) Toward real time fractal image compression using graphics hardware. Adv Vis Comput Proc Lect Notes Comput Sci 3804:723–728
13. Palazzari P, Coli M, Guglielmo L (1999) Massively parallel processing approach to fractal image compression with near-optimal coefficient quantization. J Syst Archit 45:765–779

14. Lee S, Omachi S, Aso H (2000) A parallel architecture for quadtree-based fractal image coding. In: Proceedings of 2000 international conference on parallel processing, pp 15–22
15. Hufnagl C, Uhl A (2000) Algorithms for fractal image compression on massively parallel SIMD arrays. Real-Time Imag 6:267–281
16. Bodo ZP (2004) Maximal processor utilization in parallel quadtree-based fractal image compression on MIMD Architectures. Informatica XLIX(2)
17. Haque ME, Al Kaisan A, Saniat MR (2014) GPU accelerated fractal image compression for medical imaging in parallel computing platform
18. Abdul-Malik HYS, Abdullah MZ (2018) High-speed fractal image compression featuring deep data pipelining strategy. IEEE Access 6
19. AlSaidi NMG, Ali A (2017) Towards enhancing of fractal image compression performance via block complexity. In: Annual conference on new trends in information & communications technology applications-(NTICT'2017) 7–9 Mar 2017

The Analysis of Noto Serif Balinese Font to Support Computer-assisted Transliteration to Balinese Script

Gede Indrawan, Luh Joni Erawati Dewi, I. Gede Aris Gunadi, Ketut Agustini, and I. Ketut Paramarta

Abstract This study is aimed to preserve the endangered Balinese local language knowledge through technology by analyzing Noto Serif Balinese (NSeB) font to support computer-assisted Balinese script transliteration. This research has never been done before and contributes to the future improvement of that font development for wide use on computer systems, including the smartphone. The analysis is done on Noto Balinese as the first computer-assisted transliteration to Balinese script using NSeB font that complies with Balinese Unicode. The testing result shows that NSeB font support should be enhanced by providing wider glyph for a certain appended form or *gantungan*, providing higher glyph for *tedung*, repositioning glyphs for certain clusters of vowel-sign and consonant-sign to avoid overlap displaying, and providing correction for a certain combination of syllable and long vowel. In the future, the accompanying algorithm of this study should be improved by taking care of the line breaking mechanism and by enriching its special words repository.

Keywords Balinese script · Noto Serif Balinese · Transliteration

G. Indrawan (✉) · L. J. E. Dewi · I. Gede Aris Gunadi · K. Agustini · I. Ketut Paramarta
Universitas Pendidikan Ganesha, Singaraja, Bali 81116, Indonesia
e-mail: gindrawan@undiksha.ac.id

L. J. E. Dewi
e-mail: joni.erawati@undiksha.ac.id

I. Gede Aris Gunadi
e-mail: igedearisgunadi@undiksha.ac.id

K. Agustini
e-mail: ketutagustini@undiksha.ac.id

I. Ketut Paramarta
e-mail: ketut.paramarta@undiksha.ac.id

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_54

1 Introduction

Balinese script, like other local ancient scripts in the world, is facing the extinction threat [1]. Its knowledge in the area of transliteration [2, 3] is also feared to disappear. As the national cultural asset in Indonesia, the prevention effort to this matter is conducted by the Bali Province through its regulation [4–6]. Other than from the government side, multi-side preservation efforts should have a wider impact. This research supports through the technological effort by providing the analysis of Noto Serif Balinese (NSeB) font to support computer-assisted Balinese script transliteration. It has not been studied yet and contributes to future improvement of that font development [7], including mobile implementation [8, 9] for ubiquitous learning [10, 11].

Previous related works [12–22] on transliteration to Balinese script used Bali Simbar (BS) font [23], Bali Simbar Dwijendra (BSD) font [24], Noto Sans Balinese font (NSaB) [25], and NSeB font [7]. The last two fonts are dedicated Balinese Unicode fonts [26, 27] which have the advantage to be recognized by wider computer systems, including the smartphone. References [12–14] used BS font with good transliteration results based on testing from "The Balinese Alphabet" document [28]. References [15, 16] used BSD font with good transliteration results. Additional testing from the ground-truth book [29] was used by Indrawan et al. [15]. Another use of this font was on ten learned lessons from computer-based transliteration [17]. References [18, 19] used NSaB font with good transliteration results. Extensive testing [20] using [29] was also conducted for the future improvement of the algorithm that uses this font. NSeB font was used in the study of backward compatibility transliteration [21] and non-alphanumeric transliteration [22].

Several sections construct this paper. Section introduction describes behind-the-scene on the use of NSeB font to support computer-assisted transliteration to Balinese script. This section also describes some related previous research in this area. Section method covers dedicated Balinese script Unicode of NSeB font and testing mechanism for that font support. Section result and discussion analyzes testing results. Section conclusion consists of several conclusions and suggestions.

2 Method

The analysis of Noto Serif Balinese (NSeB) font support is related to the font construction design on computer-assisted transliteration to Balinese script. It involves a certain testing mechanism on the Noto Balinese web application [30] based on NSeB font that supports dedicated Balinese Unicode [26, 27]. The testing uses part of the test in the Balinese language [18] since it exposes various aspects of NSeB font support.

3 Result and Discussion

The testing result of transliteration to Balinese script on the Noto Balinese was shown in Fig. 1a. The NSeB support which is the focus of this study related to the font construction design (represented by a solid line). Certain accompanying transliteration algorithm (represented by a dashed line) is shortly described since the analysis has already been conducted [18]. NSeB font support is analyzed based on several cases, which are numbered at the transliteration result (Fig. 1a), and based on a comparison with the ground-truth transliteration result based on BSD font [15, 16, 24], as shown in Fig. 1b.

Table 1 shows several transformation steps of input. NSeB font displays the final characters on Step 12 as Balinese script (Fig. 1a). The red string with the underline represents string transformation from the previous step. The blue string with the underline represents the string that was removed to the next step. All of these transformations were handled by a transliteration algorithm [18] and are shown here as a supporting process for the result in Fig. 1a.

Related to the font construction design, several cases were discussed, i.e.,

- For case #1 (Fig. 1a), glyph of appended form or *gantungan* "wa" of the word "luir" (namely)—constructed by the glyph of basic syllable "la" + *gantungan* "wa" + vowel-sign *ulu* "i" + consonant-sign *surang* "r"—is narrower than its counterpart glyph of BSD font (Fig. 1b). It looked obvious since the preceding glyph of the basic syllable "ᬱ" (la) is wider than other basic syllable glyphs, like "ᬲ" (sa) of word "suara" (vowel)—constructed by the glyph of the basic syllable "sa" + *gantungan* "wa" + basic syllable "ra"—for the same kind of case in case #2. The same kind of case also happened at different glyph, i.e., *gantungan* "ya" (case #3) of the word "wianjana" (consonant)—constructed by the glyph of the

Fig. 1 Transliteration result: **a** Testing; **b** Ground-truth

Table 1 Transformation steps on transliteration

No.	Latin code
1	Akeh akśarane, 47, luir ipun: akśara suara, 14, akśara wianjana, 33, akśara suara punika talĕr dados pangangge suara, tur madrĕwe suara kakalih, kawāśţanin: suara hrĕswa miwah dīrgha
2	akeh akśarane, 47, luir ipun: akśara suara, 14, akśara wianjana, 33, akśara suara punika talĕr dados pangangge suara, tur madrĕwe suara kakalih, kawāśţanin: suara hrĕswa miwah dīrgha
3	akeh \akßarane, 47, lwir ipun: \akßara swara, 14, \akßara wyanyjana, 33, \akßara swara punika talær dados pangangge swara, tur madræwe swara kakalih, kawāßttanin: swara h ræswa miwah di\nrgha
4	akeh \akßarane, ᭔᭗, lwir ipun: \akßara swara, ᭑᭔, \akßara wyanyjana, ᭓᭓, \akßara swara punika talær dados pangangge swara, tur madræwe swara kakalih, kawāßttanin: swara h ræswa miwah di\nrgha
5	akeh ᬅkßarane, ᭔᭗, lwir ipun: ᬅkßara swara, ᭑᭔, ᬅkßara wyanyjana, ᭓᭓, ᬅkßara swara punika talær dados pangang≃ge swara, tur madræwe swara kakalih, kawāßttanin: swara h ræswa miwah diᬀrgha
6	hakeh ᬅkßarane, ᭔᭗, lwir hipun: ᬅkßara swara, ᭑᭔, ᬅkßara wyanyjana, ᭓᭓, ᬅkßara swara punika talær dados pangang ~ ge swara, tur madræwe swara kakalih, kawāßttanin: swara h ræswa miwah diᬀrgha

(continued)

Table 1 (continued)

No.	Latin code
7	ᬳ ᭄ aᬓ ᭄ eᬳ ᭄ᬅ ᬓ ᭄ ᬱ ᭄ aᬭ ᭄ aᬦ ᭄ e, ᭔ ᭗, ᬮ ᭄ ᬯ ᭄ iᬭ ᭄ ᬳ ᭄ iᬧ ᭄ uᬦ ᭄: ᬅ ᬓ ᭄ ᬱ ᭄ aᬭ ᭄ a ᬲ ᭄ ᬯ ᭄ aᬭ ᭄ a, ᭑ ᭔, ᬅ ᬓ ᭄ ᬱ ᭄ aᬭ ᭄ a ᬯ ᭄ ᬬ ᭄ aᬜ ᭄ ᬚ ᭄ aᬦ ᭄ a, ᭓ ᭓, ᬅ ᬓ ᭄ ᬱ ᭄ aᬭ ᭄ a ᬲ ᭄ ᬯ ᭄ aᬭ ᭄ a ᬧ ᭄ uᬦ ᭄ iᬓ ᭄ a ᬢ ᭄ aᬮ ᭄ æᬭ ᭄ᬤ ᭄ aᬤ ᭄ oᬲ ᭄ᬧ ᭄ aᬗ ᭄ aᬗ ᭄~ ᬕ ᭄ e ᬲ ᭄ ᬯ ᭄ aᬭ ᭄ a, ᬢ ᭄ uᬭ ᭄ᬫ ᭄ aᬤ ᭄ ᬭ ᭄ æᬯ ᭄ e ᬲ ᭄ ᬯ ᭄ aᬭ ᭄ a ᬓ ᭄ aᬓ ᭄ aᬮ ᭄ iᬳ ᭄, ᬓ ᭄ aᬯ ᭄ āᬱ ᭄ ᬝ ᭄ aᬦ ᭄ iᬦ ᭄ : ᬲ ᭄ ᬯ ᭄ aᬭ ᭄ a ᬳ ᭄ᬭ ᭄ æᬲ ᭄ ᬯ ᭄ a ᬫ ᭄ iᬯ ᭄ aᬳ ᭄ᬤ ᭄ iᬀ ᬭ ᭄ ᬖ ᭄ a
8	ᬳ ᭄ aᬓ ᭄ eᬳ ᭄ ᬅ ᬓ ᭄ ᬱ ᭄ aᬭ ᭄ aᬦ ᭄ e,᭔ ᭗, ᬮ ᭄ ᬯ ᭄ iᬭ ᭄ ᬳ ᭄ iᬧ ᭄ uᬦ ᭄ :ᬅ ᬓ ᭄ ᬱ ᭄ aᬭ ᭄ aᬲ ᭄ ᬯ ᭄ aᬭ ᭄ a,᭑ ᭔, ᬅ ᬓ ᭄ ᬱ ᭄ aᬭ ᭄ aᬯ ᭄ ᬬ ᭄ aᬜ ᭄ ᬚ ᭄ aᬦ ᭄ a,᭓ ᭓, ᬅ ᬓ ᭄ ᬱ ᭄ aᬭ ᭄ aᬲ ᭄ ᬯ ᭄ aᬭ ᭄ aᬧ ᭄ uᬦ ᭄ iᬓ ᭄ aᬢ ᭄ aᬮ ᭄ æᬭ ᭄ ᬤ ᭄ aᬤ ᭄ oᬲ ᭄ ᬧ ᭄ aᬗ ᭄ aᬗ ᭄ ~ ᬕ ᭄ eᬲ ᭄ ᬯ ᭄ aᬭ ᭄ a,ᬢ ᭄ uᬭ ᭄ ᬫ ᭄ aᬤ ᭄ ᬭ ᭄ æᬯ ᭄ eᬲ ᭄ ᬯ ᭄ aᬭ ᭄ aᬓ ᭄ aᬓ ᭄ aᬮ ᭄ iᬳ ᭄, ᬓ ᭄ aᬯ ᭄ āᬱ ᭄ ᬝ ᭄ aᬦ ᭄ iᬦ ᭄ :ᬲ ᭄ ᬯ ᭄ aᬭ ᭄ aᬳ ᭄ ᬭ ᭄ æᬲ ᭄ ᬯ ᭄ aᬫ ᭄ iᬯ ᭄ aᬳ ᭄ ᬤ ᭄ iᬀ ᬭ ᭄ ᬖ ᭄ a

(continued)

Table 1 (continued)

No.	Latin code
9	ᬳ ᬓ ᬾ ᬳ ᭄ ᬅ ᬓ ᭄ ᬱ ᬭ ᬦ ᬾ, ᭔ ᭗, ᬮ ᭄ ᬯ ᬶ ᬭ ᭄ ᬳ ᬶ ᬧ ᬸ ᬦ ᭄ :ᬅ ᬓ ᭄ ᬱ ᬭ ᬲ ᭄ ᬯ ᬭ, ᭑ ᭔, ᬅ ᬓ ᭄ ᬱ ᬭ ᬯ ᭄ ᬬ ᬜ ᭄ ᬚ ᬦ, ᭓ ᭓, ᬅ ᬓ ᭄ ᬱ ᬭ ᬲ ᭄ ᬯ ᬭ ᬧ ᬸ ᬦ ᬶ ᬓ ᬢ ᬮ ᭂ ᬭ ᭄ ᬤ ᬤ ᭀ ᬲ ᭄ ᬧ ᬗ ᬗ ᭄ ~ ᬕ ᬾ ᬲ ᭄ ᬯ ᬭ, ᬢ ᬸ ᬭ ᭄ ᬫ ᬤ ᭄ ᬭ ᭂ ᬯ ᬾ ᬲ ᭄ ᬯ ᬭ ᬓ ᬓ ᬮ ᬶ ᬳ ᭄, ᬓ ᬯ ᬵ ᬱ ᭄ ᬝ ᬦ ᬶ ᬦ ᭄ :ᬲ ᭄ ᬯ ᬭ ᬳ ᭄ ᬭ ᭂ ᬲ ᭄ ᬯ ᬫ ᬶ ᬯ ᬳ ᭄ ᬤ ᬶ ᬀ ᬭ ᭄ ᬖ
10	ᬳ ᬓ ᬾ ᬳ ᭄ ᬅ ᬓ ᭄ ᬱ ᬭ ᬦ ᬾ, ᭔ ᭗, ᬮ ᭄ ᬯ ᬶ ᬭ ᭄ ᬳ ᬶ ᬧ ᬸ ᬦ ᭄ :ᬅ ᬓ ᭄ ᬱ ᬭ ᬲ ᭄ ᬯ ᬭ, ᭑ ᭔, ᬅ ᬓ ᭄ ᬱ ᬭ ᬯ ᭄ ᬬ ᬜ ᭄ ᬚ ᬦ, ᭓ ᭓, ᬅ ᬓ ᭄ ᬱ ᬭ ᬲ ᭄ ᬯ ᬭ ᬧ ᬸ ᬦ ᬶ ᬓ ᬢ ᬍ ᬭ ᭄ ᬤ ᬤ ᭀ ᬲ ᭄ ᬧ ᬗ ᬗ ᭄ ~ ᬕ ᬾ ᬲ ᭄ ᬯ ᬭ, ᬢ ᬸ ᬭ ᭄ ᬫ ᬤ ᬺ ᬯ ᬾ ᬲ ᭄ ᬯ ᬭ ᬓ ᬓ ᬮ ᬶ ᬳ ᭄, ᬓ ᬯ ᬵ ᬱ ᭄ ᬝ ᬦ ᬶ ᬦ ᭄ :ᬲ ᭄ ᬯ ᬭ ᬳ ᬺ ᬲ ᭄ ᬯ ᬫ ᬶ ᬯ ᬳ ᭄ ᬤ ᬶ ᬀ ᬭ ᭄ ᬖ

(continued)

Table 1 (continued)

No.	Latin code
11	ᬳ ᬓ ᬾ ᬄ ᬅ ᬓ ᭄ ᬱ ᬭ ᬦ ᬾ, ᭔ ᭗, ᬮ ᭄ ᬯ ᬶ ᬃ ᬳ ᬶ ᬧ ᬸ ᬦ ᭄ :ᬅ ᬓ ᭄ ᬱ ᬭ ᬲ ᭄ ᬯ ᬭ, ᭑ ᭔, ᬅ ᬓ ᭄ ᬱ ᬭ ᬯ ᭄ ᬬ ᬜ ᭄ ᬚ ᬦ, ᭓ ᭓, ᬅ ᬓ ᭄ ᬱ ᬭ ᬲ ᭄ ᬯ ᬭ ᬧ ᬸ ᬦ ᬶ ᬓ ᬢ ᬍ ᬃ ᬤ ᬤ ᭀ ᬲ ᭄ ᬧ ᬗ ᬗ ᭄ ᬕ ᬾ ᬲ ᭄ ᬯ ᬭ, ᬢ ᬸ ᬃ ᬫ ᬤ ᬺ ᬯ ᬾ ᬲ ᭄ ᬯ ᬭ ᬓ ᬓ ᬮ ᬶ ᬄ, ᬓ ᬯ ᬵ ᬱ ᭄ ᬝ ᬦ ᬶ ᬦ ᭄ :ᬲ ᭄ ᬯ ᬭ ᬳ ᬺ ᬲ ᭄ ᬯ ᬫ ᬶ ᬯ ᬄ ᬤ ᬶ ᬀ ᬃ ᬖ
12	ᬳ ᬓ ᬾ ᬄ ᬅ ᬓ ᭄ ᬱ ᬭ ᬦ ᬾ ᭞ ᭔ ᭗ ᭞ ᬮ ᭄ ᬯ ᬶ ᬃ ᬳ ᬶ ᬧ ᬸ ᬦ ᭄ ᭝ ᬅ ᬓ ᭄ ᬱ ᬭ ᬲ ᭄ ᬯ ᬭ ᭞ ᭑ ᭔ ᭞ ᬅ ᬓ ᭄ ᬱ ᬭ ᬯ ᭄ ᬬ ᬜ ᭄ ᬚ ᬦ ᭞ ᭓ ᭓ ᭞ ᬅ ᬓ ᭄ ᬱ ᬭ ᬲ ᭄ ᬯ ᬭ ᬧ ᬸ ᬦ ᬶ ᬓ ᬢ ᬍ ᬃ ᬤ ᬤ ᭀ ᬲ ᭄ ᬧ ᬗ ᬗ ᭄ ᬕ ᬾ ᬲ ᭄ ᬯ ᬭ ᭞ ᬢ ᬸ ᬃ ᬫ ᬤ ᬺ ᬯ ᬾ ᬲ ᭄ ᬯ ᬭ ᬓ ᬓ ᬮ ᬶ ᬄ ᭞ ᬓ ᬯ ᬵ ᬱ ᭄ ᬝ ᬦ ᬶ ᬦ ᭄ ᭝ ᬲ ᭄ ᬯ ᬭ ᬳ ᬺ ᬲ ᭄ ᬯ ᬫ ᬶ ᬯ ᬄ ᬤ ᬶ ᬀ ᬃ ᬖ ᭟

basic syllable “wa” + *gantungan* “ya” + basic syllable “nya” + *gantungan* “ja” + basic syllable “na”, number “3” (case #4) of the word “33” (constructed by two glyphs of the number “3”), and *gantungan* “rě” (case #6) of the word “hrěswa” (kind of vowel)—constructed by the glyph of the basic syllable “ha” + *gantungan* “rě” + basic syllable “sa” + *gantungan* “wa”.

- For case #5 (Fig. 1a), glyph of vowel-sign *tedung* “ ◌ᬵ” (in the form of *ligature*) of word “dados” (become)—constructed by the glyph of the basic syllable “da” + vowel-sign *taling* “e” + basic syllable “da” + vowel-sign *tedung* “o” + basic syllable “sa” + consonant-sign *adeg-adeg*—is shorter than its counterpart glyph of BSD font (Fig. 1b). Normally, its height is identical to the height of the preceding base character “ ᬤ” (da). The glyph of NSeB’s Balinese Musical Symbol Dong “ ᭡” is more appropriate in this case [31].
- For case #7 (Fig. 1a), there are overlapping on glyph displaying between vowel-sign *ulu-sari* “ī” and consonant-sign *surang* “r” of word “dīrgha” (kind of vowel)—constructed by the glyph of basic syllable “da” + vowel-sign *ulu-sari* “ī” + consonant-sign *surang* “r” + basic syllable “gha”. This case should not have happened, as shown by its counterpart glyph displaying of BSD font (Fig. 1b), where “ī” and “r” were displayed on the left side and right side, respectively. For the same kind of case, proper displaying of NSeB has already been conducted between “i” and “r” in case #1.

Related to the accompanying algorithm to show other aspects of Noto Balinese, several cases were discussed, i.e.,

- For case #8, line breaking mechanism has not become a concern yet [14], as it involves two rules of thumb, i.e., (1) no line breaking is allowed between syllable and any of its signs; and (2) no line breaking is allowed just before a colon, comma, or full stop.
- Special transformations of string (Table 1) depend on exceptions to the rule. A consonants combination example in case #9 [32] on word “wianjana”, “nj” was converted to “nyj” for correct transliteration. In a recent development, this kind of transformation is maintained by using a database [21].

Another aspect related to NSeB font support was revealed in [21], i.e., on the incorrect display of Balinese script of “ghā” as a combination of syllable *ga gora* “gha” and long vowel (represented by vowel-sign *tedung* “ ◌ᬵ”). On the Noto Balinese, “ghā” was transliterated as “ ᬕᬵ” (belong to a combination of syllable “ga” and long vowel) rather than the right “ ᬖᬵ”. It is considered as a flaw in NSeB font construction design since the algorithm treatment for this kind of combination gives the correct display of Balinese script. For example, “cā dhā hā kā mā śā tā ţā wā yā” [28] was transliterated correctly as “ ᬘᬵ ᬥᬵ ᬳᬵ ᬓᬵ ᬫᬵ ᬰᬵ ᬢᬵ ᬝᬵ ᬯᬵ ᬬᬵ”.

4 Conclusion

Noto Serif Balinese (NSeB) font should be improved in the future development to support computer-assisted Balinese script transliteration. The future work of this study is in the accompanying algorithm of this computer-assisted transliteration application that uses NSeB font. The improvement can be conducted by taking care of the line breaking mechanism and by enriching its special words repository from the Balinese script dictionary.

Acknowledgements This work is part of a research grant funded by the Indonesian Ministry of Education, Culture, Research, and Technology.

References

1. Indrawan G, Paramarta IK, Agustini K, Sariyasa S (2018) Latin-to-Balinese script transliteration method on mobile application: a comparison. Indones J Electr Eng Comput Sci 10(3):1331
2. Karimi S, Scholer F, Turpin A (2011) Machine transliteration survey. ACM Comput Surv 43(3):1–46
3. Kaur K, Singh P (2014) Review of machine transliteration techniques. Int J Comput Appl 107(20):13–16
4. Bali Governor Circular Letter No. 3172 (2019) https://jdih.baliprov.go.id/produk-hukum/peraturan/abstrak/24741. Last accessed 2021/08/17
5. Bali Government Regulation No. 3 (1992) https://bphn.go.id/data/documents/perda_3_1992.pdf. Last accessed 2021/08/17
6. Bali Governor Regulation No. 80 (2018) https://jdih.baliprov.go.id/produk-hukum/peraturan/abstrak/24665. Last accessed 2021/08/17
7. Google Noto Serif Balinese. https://github.com/googlefonts/noto-fonts/blob/master/unhinted/ttf/NotoSerifBalinese/NotoSerifBalinese-Regular.ttf. Last accessed 2021/08/17
8. Griffith C (2017) Mobile app development with ionic 2: cross-platform apps with Ionic, Angular, and Cordova. O'Reilly, CA
9. Cha S, Yun Y (2013) Smartphone application development using HTML5-based cross-platform framework
10. Hwang GJ et al (2008) Criteria, strategies and research issues of context-aware ubiquitous learning. J Educ Technol Soc 11(2)
11. Ogata H et al (2009) LORAMS: linking physical objects and videos for capturing and sharing learning experiences towards ubiquitous learning. Int J Mob Learn Org 3(4):337–350
12. Indrawan G, Sariyasa, Paramarta IK (2019) A new method of Latin-to-Balinese script transliteration based on Bali Simbar font. In: The 4th international conference on informatics and computing
13. Crisnapati PN et al (2019) Pasang Aksara Bot: A Balinese script writing robot using finite state automata transliteration method. J Phys: Conf Ser 1175(1):012108
14. Indrawan G, Setemen K, Sutaya IW, Paramarta IK (2020) Handling of line breaking on Latin-to-Balinese script transliteration web application as part of Balinese language ubiquitous learning. In: The 6th international conference on science in information technology
15. Indrawan G et al (2020) An improved algorithm and accuracy analysis testing cases of Latin-to-Balinese script transliteration method based on Bali Simbar Dwijendra Font. Test Eng Manage 83:7676–7683

16. Indrawan G, Dantes GR, Aryanto KYE, Paramarta IK (2020) Handling of mathematical expression on Latin-to-Balinese script transliteration method on mobile computing. In: The 5th international conference on informatics and computing
17. Indrawan G, Gunadi IGA, Gitakarma MS, Paramarta IK (2021) Latin to Balinese script transliteration: lessons learned from the computer-based implementation. In: The 4th international conference on software engineering and information management
18. Indrawan G, Paramarta IK, Agustini K (2019) A new method of Latin-to-Balinese script transliteration based on Noto Sans Balinese font and dictionary data structure. In: The 2nd international conference on software engineering and information management
19. Indrawan G, Puspita NNH, Paramarta IK, Sariyasa S (2018) LBtrans-Bot: a Latin-to-Balinese script transliteration robotic system based on Noto Sans Balinese Font. Indones J Electr Eng Comput Sci 12(3):1247
20. Loekito LH, Indrawan G, Sariyasa S, Paramarta IK (2020) Error analysis of Latin-to-Balinese Script transliteration method based on Noto Sans Balinese Font. In: The 3rd international conference on innovative research across disciplines
21. Indrawan G, Paramarta IK, Nurhayata IG, Sariyasa S (2021) A method to accommodate backward compatibility on the learning application-based transliteration to the Balinese script. Int J Adv Comput Sci Appl 12(6):280–286
22. Indrawan G, Sutaya IW, Ariawan KU, Gitakarma MS, Nurhayata IG, Paramarta IK (in press) A method for non-alphanumeric text processing on transliteration to the Balinese Script. ICIC Exp Lett
23. Bali Simbar, http://www.babadbali.com/aksarabali/balisimbar.htm. Last accessed 2021/08/17
24. Suatjana IM (2009) Bali Simbar Dwijendra. Yayasan DwijendraDenpasar
25. The Unicode Consortium (2020) The Unicode Standard Version 13.0—Core Specification. The Unicode Consortium
26. Balinese Unicode Table, http://unicode.org/charts/PDF/U1B00.pdf. Last accessed 2021/08/17
27. Google Noto Fonts, https://www.google.com/get/noto/#sans-bali. Last accessed 2021/08/17
28. The Balinese Alphabet, http://www.babadbali.com/aksarabali/alphabet.htm. Last accessed 2021/08/17
29. Anom IGK et al (2009) Balinese-Indonesian Dictionary with its Latin and Balinese Script [Kamus Bali-Indonesia Beraksara Latin dan Bali]. Bali Province, Denpasar
30. Groner L, Manricks G (2015) JavaScript regular expressions. Packt Publishing, Birmingham
31. Various Tedung Defects in Noto Serif Balinese Phase III, https://github.com/googlefonts/noto-fonts/issues/1335. Last accessed 21 July 2020
32. Medra IN et al (1997) The guidance of Balinese Script Construction [Pedoman Pasang Aksara Bali]. Department of Culture, Bali Province, Denpasar

Email Spam Detection Using Multilayer Perceptron Algorithm in Deep Learning Model

Senthil Murugan Tamilarasan, Muthyala Hithasri, and Kamakshi Pille

Abstract Email spam detection is a filtering process which identifies either it is spam or not. It also removes the unsolicited data present in the user's email inbox. Certain type of spam mails contains malware which misuse the users' data. Hence, we need to identify spam mails and take necessary actions. Many machine learning algorithms have proposed for differentiate spam mails from normal mails. Tokenization of emails between length and frequency is one of the techniques. It helps to split the raw emails into tokens known as small words. After tokenization, tokenized count has taken into consideration for process the emails. Based on that spam emails to be identified which are present in the dataset of spam and ham emails. To extracting these features, term frequency—inverse document frequency (TF-IDF) has used to train the model. In this method, multilayer perceptron deep learning algorithm is applied to compute the model. It has two layers. When input is given to the perceptron, the input is multiplied by the hidden layers, and it holds the activation function such as sigmoid activation with regularization function. For the better optimization, the model uses the Adam optimizer with gradient descent for fastest optimization. The network learns the model. The learning rate is set to true. While computing the model, it goes in the forward direction to train the model and comeback again (backpropagation). This process will be repeated. Going to the forward direction and comes back, then again, maintaining forward approach is called one epoch. The epoch rate has computed in the model. In the comparison between multilayer perceptron algorithm and machine learning algorithms such as support vector machine (SVM), random forest, and XGBoost, the deep learning algorithm produces 99% of accuracy on precision, recall, and F-measure and holds less computation time. Hence, the results prove that deep learning algorithm performs better than machine learning algorithms.

Keywords Email spam detection · Deep learning model · Multilayer perceptron

S. M. Tamilarasan · M. Hithasri (✉) · K. Pille
Kakatiya Institute of Technology and Science, Warangal, Telangana, India
e-mail: hithasri.muthyala@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_55

1 Introduction

Spam mails may be sending to user's inbox directly. Unfortunately, user may open the mail or click some links which is in that mail. It may cause for stoles the users' personal data by frauds. In these days, apart from email, certain type of spam is increasing through text message services. Such type of spamming occurs in various categories in many ways. In this paper, we are going to prepare spam filter and identifying them either spam or not spam. Previously, various types of machine learning algorithms were used for detection of spam mails. But, we are going to identify spam through deep learning model with the techniques of multilayer perceptron. We expect to achieve 99% accuracy in performance, f1-score, and recall. The detection takes place by many of the processing techniques like importing of libraries and preparing of Enron dataset which easily distinguishes between ham and spam, and the filter consists of many layers, and few types of dictionaries are used.

Khan et al. [1] proposed TF-IDF model is created for extracting the features from tokenizing the input elements present in the data. Keras software tool acts as front end, and TensorFlow act as backend. Layer1 (l1) and layer2 (l2) consist of neurons with linear ReLU, with sigmoid activation function. Sigmoid is an activation function present in the deep neural networks. Regularization is a technique reduces the complexity in neural networks. Here, regularization function is used to solve more complex problems. Dropout is taken with probability rate is used to prevent the problem of over fitting in l1 and l2. Loss function is implemented on binary cross-entropy and for fastest optimization in gradient descent estimation.

Gradient descent is an optimization technique. Adam optimizer is used to better optimization and data shuffling which helps to improve the performance of the model. This model has chosen batch size and learning rate. If the learning rate increases when compute the model, it takes more time which is present in the deep learning. Features are extracted automatically in the deep learning model. Deep learning takes more time when compared to the machine learning. But, the problem with the machine learning is those features that are not trained automatically. We have to train the model. We should give some type of inputs. Machine learning algorithms are hands-crafted one, while deep learning model takes automatically trained the model.

2 Literature Survey

In machine learning techniques, the following methods are used to find the spam mails. Such as support vector machine, random forest, decision tree, and Naive Bayes. Also, those techniques have been compared with bio-inspired metaheuristic algorithms. Some genetic algorithm and particle swarm optimization algorithm are also used to detect spam mails. These algorithms are implemented using seven different types of datasets and different tools such as WEKA, scikit-learn, Keras, TensorFlow. Srinivasan et al. [2] used Python platforms on Jupiter Notebook which is an

open-source tool and also used similar to Spyder ide. In Spyder, the programmer can capable to include multiple panels such as console where the output can be visible and variable explorer which assignment of variables. Online platforms such as Google Collaboratory, Kaggle desktop top-based platforms are also used.

The Enron-Spam dataset contains the following six datasets. SpamAssassin dataset converts folder into text files. Gibson et al. [3] introduced a Ling-Spam dataset contains ten parts where dataset is not preprocessed. PUA is a type of numerical dataset that has different types of mails. Dataset has been splitted into different ratios like 60:40, 70:30, 75:25 and 80:20 for training and testing respectively. Scikit-learn is the library they had used to measure the performance. Recall, precision, f1-score, accuracy had been calculated using confusion matrix, on every dataset. Some of the tuning parameters has been applied by Gangavarapu et al. [4] for detection of emails like SGD, MNB, MLP, DT, and RF. A particle swarm optimization (PSO) which is bio-inspired algorithms considers the best evaluation position and global position for usage of machine learning models. It is used in PSO for the purpose of feature selection. Number of particles also used as input for this model and it is considered from library.

Genetic algorithm iterates through a fitness function between two individuals and produces offspring. They have been tested the models with the 50,000 emails. Awad and Foqaha [5] tested on the numerical corpus, restrictions occurred in the extracting the features. While they implemented on alphabetical corpus, it had given the best results in extracting the features and predicting outcome. WEKA acts as a black box that ran the datasets on 14 different classification algorithms. It produces best results in the top algorithms like multinomial Naive Bayes, support vector machine, random forest, and decision tree. All these algorithms have been tested and experimented with scikit-learn library and its related modules by Hussain et al. [6].

Genetic algorithm works better on the text-based datasets as well as numerical datasets than PSO. PSO worked efficiently with the multinomial Naive Bayes and stochastic gradient descent. Whereas, genetic algorithms worked well for decision tree and random forest. On overall results, Naive Bayes algorithm works better for detection of spam detection of emails, and multinomial Naive Bayes performed highest accuracy of 98% with genetic algorithm optimization for 80:20 split of train and test on spam assassin dataset.

3 Machine Learning Models

Three models like random forest, support vector machine, and XGBoost extracting the TF-IDF feature have been tested for compare the results. Many researchers have done on the machine learning models like Naive Bayes support vector machine, decision tree, XGBoost, random forest for spam detection.

3.1 Support Vector Machine (SVM)

In this method, train the classifier to take learning features and learn how they relate to the training of species and iterates. It has done over the linear data and sample data to optimize linearly in a hyper-plane stochastic gradient descent.

3.2 Random Forest

Random forest algorithm is used for both classification and regression model implement on scikit-learn library. It uses on termination criteria means; it goes with deep depth for computation process. It takes more time. Evaluation goes on random forest model and selects the higher number of classes as prediction. Random forest is an algorithm depends on the decision tree. If tree is complex, then go for the random forest algorithm where decision tree goes in an top-down approach manner.

3.3 XGBoost

XGBoost is known as extreme gradient boosting; it implements on the gradient-boosted decision trees designed for speed and performance. It can also be used for classification and regression model in training the model. XGBoost is taken as XGBoost and fit model.

4 Different Steps in Deep Learning

This section describes the step-by-step processing of deep learning techniques.

4.1 Preprocessing the Data

Preprocessing results in creating an Enron dataset and generate pandas' data frame. Some analyses are applied on the data like tokenizing. First step of model is importing the libraries and preparing an Enron dataset into the framework, load, and extract the Enron data spam in panda's framework. After loading the Enron data, data frame is applied for tokenized text, tokenized count, and language.

4.2 Preparation of Training and Test Data

Split the dataset for training and validation. We are going to take approximately 10,000 emails for testing, and remaining emails take place for build the model. The data are going to shuffle in the data frame.

4.3 Feature Generation

In the process of feature generation, unsupervised learning algorithm is converted to the supervised learning algorithm using the labeled Enron-Spam dataset. TF-IDF models are used for counting number of spam words, stop words, repeating words held in the dataset for training the models. Importing Keras library in TensorFlow acts as back end and prepare the function using TF-IDF feature models for generation of input. This model has taken totally 34 s for building the model.

4.4 Split and Validation of Data

This step is useful in order to prevent the model from overfitting. Here, split the data for training and testing. Collecting a sample of data remained back from the training model results in validation phase and estimating it in tuning the models in hyper-parameters present in the data. Hence, the entire dataset has been splitted into 85:15 ratio for training and testing respectively. We are importing test, train, and split for validation, and model selection is used from the Sklearn library.

4.5 Model Building

In this model building, it takes two input layers as layer one consists of linear ReLU, and layer two consists of linear ReLU with sigmoid function which is an activation function used to predict the probability of the model. This sigmoid activation function is mostly used in the multilayer neural networks. It predicts in shape ‘s’ which is a nonlinear activation function. Layer one consists of 512 neurons, and layer two consists of 256 neurons with sigmoid activation function to avoid the problem of overfitting. A regularization function is used dropout rate and also added for it with 0.5 probabilities in l1 and l2. Loss function is used for back propagating the model in a sequence manner. Loss function is applied on the binary cross-entropy to get optimization for faster estimation.

Adam optimizer is using gradient descent estimation to perform accurately on parameters and to minimize the cost function. Data shuffling meant that shuffle the

data within an attribute maintained at true and batch size is taken as 64. The network has initialized with 0.001 learning rate, and check points were kept spam detection of spam takes place.

5 Results

Deep learning model with multilayer perceptron does very well on the test data. The results from other models are close but more. We tried this approach over multiple language emails, and deep learning model is very consistent with the performance. XGBoost also does very well.

In confusion matrix, true-positive rate means when the model is predicted true and output also predicted as true. True-negative rate means when the model is predicted positive and output is predicted as negative. False-positive rate means when the model is false and also output predicted as negative and false positive rate means when the model is false and output is predicted as negative (Table 1).

Table 2 shows the classified outputs in the following rates are true positive and false negative because they are results in the same as present in the input. These rates are used in the data-related topics for finding the accuracy how much actively our preferred type of model is performing related to our approach on the confusion matrix; these types of values are predicted; confusion matrix is divided into four blocks in which each size has a rate such as true positive, true negative, false positive, and false negative using this confusion matrix; only, the predicted output is generated perfectly by using the model extraction. F1-score is also called as f-measure where f is a randomly chosen. Precision finds the accurateness of values the relevant instances which are retrieved among the instances, and recall is also called as the sensitivity

Table 1 Confusion matrix

Model	False negative	True positive	True negative	False positive
Random forest	732	5187	3743	338
SVM	540	5379	4026	55
XGBoost	479	5440	3398	683
Deep learning	46	5873	4037	44

Table 2 Comparison results based on precision, recall, F1-score, and samples

Model	Precision	Recall	F1-score	Total samples
Random forest	0.887624	0.896754	0.890721	10,000
SVM	0.935807	0.947646	0.939390	10,000
XGBoost	0.882452	0.875857	0.878744	10,000
Deep learning	0.990649	0.990723	0.990686	10,000

used find the relevant instances that are retrieved actual instances.

$$\text{Precision} = (\text{True positive})/(\text{True positive} + \text{false positive}) \quad (1)$$

$$\text{Recall} = (\text{True positive})/(\text{True positive} + \text{false negative}) \quad (2)$$

$$F1\ \text{score} = 2 * ((\text{Precision} * \text{recall})/(\text{Precision} + \text{recall})) \quad (3)$$

6 Conclusion

We concluded that deep learning model with the multilayer perceptron algorithm produces the accuracy of 99% in precision, recall, F1-score. The final results are compared with machine learning algorithms such as random forest, SVM, and XGBoost, and it is shown in Table 2. We considered only 1000 samples from Enron dataset for the testing and generated confusion matrix which shown in Table 1. Finally, deep learning algorithm produced more good results compared with machine learning algorithms. It may be vary when using different datasets and total number of data at the time of tuning. In future, the deep learning researchers may apply bio inspired based algorithm for improving the results in this model.

References

1. Khan WZ, Khan MK, Muhaya FT, Aalsalem MY, Chao HC (2015) A comprehensive study of email spam botnet detection. IEEE Commun Surv Tutorials 17(4)
2. Srinivasan S, Ravi V, Alazab M, Ketha S, Al-Zoubi AM, Kotti Padannayil S (2021) Spam emails detection based on distributed word embedding with deep learning. In: Maleh Y, Shojafar M, Alazab M, Baddi Y (eds) Machine intelligence and big data analytics for cybersecurity applications. Studies in computational intelligence, vol 919. Springer, Cham
3. Gibson S, Issac B, Zhang L, Jacob SM (2020) Detecting spam email with machine learning optimized with bio-inspired metaheuristic algorithms. IEEE Access 8:187914–187932. https://doi.org/10.1109/ACCESS.2020.3030751
4. Gangavarapu T, Jaidhar CD, Chanduka B (2020) Applicability of machine learning in spam and phishing email filtering: review and approaches. Artif Intell Rev 53:5019–5081
5. Awad M, Foqaha M (2016) Email spam classification using hybrid approach of RBF neural network and particle swarm optimization. Int J Netw Secur Its Appl 8(4):17–28
6. Hussain N, Mirza HT, Hussain I, Iqbal F, Memon I (2020) Spam review detection using linguistic and spammer behavioural methods vol 8. IEEE Access 2020

Robust Feature Extraction and Recognition Model for Automatic Speech Recognition System on News Report Dataset

Sunanda Mendiratta, Neelam Turk, and Dipali Bansal

Abstract Information processing has become ubiquitous. The process of deriving speech from transcription is known as automatic speech recognition systems. In recent days, most of the real-time applications such as home computer systems, mobile telephones, and various public and private telephony services have been deployed with automatic speech recognition (ASR) systems. Inspired by commercial speech recognition technologies, the study on automatic speech recognition (ASR) systems has developed an immense interest among the researchers. This paper is an enhancement of convolution neural networks (CNNs) via a robust feature extraction model and intelligent recognition systems. First, the news report dataset is collected from a public repository. The collected dataset is subjective to different noises that are preprocessed by min–max normalization. The normalization technique linearly transforms the data into an understandable form. Then, the best sequence of words, corresponding to the audio based on the acoustic and language model, undergoes feature extraction using Mel-frequency Cepstral Coefficients (MFCCs). The transformed features are then fed into convolutional neural networks. Hidden layers perform limited iterations to get robust recognition systems. Experimental results have proved better accuracy of 96.17% than existing ANN.

Keywords Speech recognition · Text · Mel features · Recognition accuracy · Convolutional neural networks

1 Introduction

In recent days, automatic speech recognition is being studied and has benefitted the application users. A distant microphone system is an instance that is widespread

S. Mendiratta (✉) · N. Turk
Department of Electronics Engineering, JCBoseUST, Faridabad, India
e-mail: sunanda.mendiratta@gmail.com

D. Bansal
GB Pant Okhla Campus 1, Delhi Skill & Entrepreneurship University, Delhi, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_56

among household users. In today's living system, a high level of background noise is encountered that develops numerous challenges for recognition systems [1]. Though the disturbing noise is recognized, in some scenarios, the performance of intelligent recognition systems is not enhanced. Automatically, it interrupts the speech signal, and this noise masking causes phonetic information loss. Thus, the recognition error rate is high due to the misinterpretation of the speech signal. In a few cases, misinterpretation due to the sensitivity of the speakers degrades the effects of speech production and noise masking. This is being resolved by the automatic speech recognition system that aimed for proper handling of speech and its background noise [2]. It generally converts the speech signal into text. Since it is a developing field; the research on man–machine interaction will provide novel insights. Speech signals are one of the complex signals which make the automatic speech recognition systems (ASRS) field, an innovative process.

The robust feature extraction process [3] determines the success rate of intelligent speech recognition systems. The ideal representation of a feature vector should carry an error-free speech signal. Hence, the examination of robust feature extraction has been major research in the ASR field. Though there are several robust feature extraction techniques available from past decades, the common issues like bandwidth and the number of filters matter the most. Generally, types of ASR systems are categorized based on the number of speakers, utterance nature, vocabulary size, and bandwidth. By understanding, the production of speech sounds and the source of variability determines the level of ASRS functionalities [4]. Unlike text, phonetics (or) words do not have boundaries due to the incessant flow of speech. This simple sample defines the difficulty level of speech interpretation; an utterance of six sheep may be sick sheep, even in the absence of noise. Also, in a printed text, multiple occurrences of a letter appear the same. In contrast, spectral and temporal characteristics of speech sound vary a lot, which depends on several factors.

When articulation errors occur consistently and, as a result, they are predictable, there exists the advantage of using ASR, even for speech that is highly unintelligible for human listeners. To our knowledge, there are not any commercial or open-source products available that would enable people in this user community to enter unrestricted text into a personal computer via automatic speech recognition [5]. The necessity of a greater amount of training data is a major concern in ASR development, especially for dysarthria speakers since speaking can be a tiring task. The rest of the paper is organized as follows: Sect. 2 presents the related work; Sect. 3 presents the proposed methodology; Sect. 4 presents the experimental analysis and finally the conclusion in Sect. 5.

2 Related Work

This section presents the existing techniques given by other researchers. The author in [6] developed a model for severe speech impairment. The training model is composed of small vocabulary and the speaker-dependent recognizer model with reduced

features extraction process. The results showed that when perplexity increases, the recognition rate increases. The analysis of redundant words is not focused on. In [7], the authors presented a Bayesian separation model for sparsity promotion and the perceptual wavelet domain. The wavelet coefficient is estimated from the discriminative information in the time–frequency plane. The model achieved a low signal-to-noise ratio with degraded performance. The sensitivity of the speech degrades the performance of the hybrid classifier. In [8], the authors discussed the silent speech recognition model for persons with laryngectomy. A set of phrases were used for training features with better running speech. Detecting the onset and offset of speech throws errors. Whisper is also a common issue that may exchange speech signals. Vector Taylor series (VTS) is used for distinguishing the background from whispering sound. Then, the model utterance defines its pseudo-whisper samples. Cepstral features were developed frame by frame to derive whisper statistics. The analysis of transcribed whispers created complexity over desired targets.

The author in [9] presented end-to-end deep learning models for speech reverberation and the acoustic modeling in which time-based DNN speech reverberation architecture to find the speech quality. Then, multi-channel microphone arrays were developed for training the DNN-based multi-conditioning. The application of nonlinear approaches still has not resolved the classical DE reverberation problem. Further, biometric recognition using HTM spatial pooler and temporal memory is performed. An array of circuit design is examined over two datasets, namely face recognition and speech recognition. Autocorrelation of each speech frame incurs design costs. Sampling rates and the delay in the chip cause poor performance. In [10], the authors studied a couple of dictionaries for exemplar radar-based speech enhancement. Initially, full-resolution frequency between speech and noises is computed and stored as corpus dictionaries. Spectrogram features of a dictionary containing the exemplar help to achieve the full-resolution frequency domain. Based on the temporal context, the computational complexity of variant noise cases decreases. Data feeding on lower-dimensional Mel features increases the mapping models.

In [11], the authors suggested a hybrid framework for developing a learning model using hidden Markov models. The log-likelihood score of each speech is computed and fed into a discriminative classifier. Finally, the deep neural network-based HMM enhanced the recognition accuracy. The intelligibility level of each uttered word operated on the choice of reference set degraded the accuracy. Imprecise and missing consonants found in these words of dysarthria speech led to overlap among the respective classes. In [12], the authors studied multi-views multi-learners approach using multi-nets ANN. The variability of dysarthria speech was analyzed by the likelihood of vocabulary words and then fed into the learner's phases, both dependent and independent paradigms. Analysis of feature vectors on associated words degrades the scalability and reliability due to incorrect classifications.

In [13], the authors presented a piece of articulatory information for ASR systems. The study on the speech inversion process is analyzed in the form of vocal tract constriction variables. Aurora 2 speech corpus was utilized for training models. Then, word recognition tasks were performed using articulatory conjunction and enhanced

the word recognition rates. Feature extraction over noises must be focused on developing a robust environment. EMG-based speech recognition using articulatory facial muscles helps tackling speaking mode variations and discrepancies between audible and inaudible speech. A spectral mapping algorithm was introduced between muscle contraction and the word utterance. The presence or absence of phonetics makes a drastic difference. Maximum entropy gain (MEG) determines the different signaling modes between audible and silent speech. Though it yields average computational complexity, the silent speech analysis must improve.

Significance of speech manifolds analyzed over cross-lingual and the multilingual by Sahraeian and Van Compernolle [14]. Manifold technique was studied for data selection and graph construction using deep neural networks. The representation of the features exhibited the language interdependency. The examined results stated better usability, whereas the analysis of inter-cross languages degrades the local structure of data. In [15], the authors studied whispered speech recognition using deep noise encoders and estimated the measures like cepstral distances, cepstral coefficients, confusion matrix, and the inverse filtering which help to classify the signals. The misclassification due to train/test scenario is of challenging tasks. In [16], the authors presented speech enhancement by improvising full-sentence correlation and clean speech recognition. The effectiveness of the speech segment and the speech utterances was estimated with predicted and unpredicted noises. The constrained maximization problem was resolved for different noises and overlapping data. Yet, the systems reduced data overlapping for small corpus.

In [17], the authors presented an energy-efficient speech extraction process in mobile head-mounted display systems. Fast speech extraction (Fast SE) algorithm is adopted for the speech selection process which degrades the low latency speech extraction. Reconfigurable matrix operation accelerator is deployed for the energy-efficient process. Configurability and debugging of each network interface communication explored higher delay time. In [18], the authors presented a long span of temporal patterns in hierarchical deep network models. Here, link quality between posterior estimates of the ASR performance was studied. In [19], the authors studied acoustic coprocessors using hidden Markov model (HMM) integrated 8-way data path of NOR flash array and the senones acoustic library. This system significantly reduced the error rate of 15.4%, but the maintenance of an acoustic library is difficult. In [20], the analysis of regularized speaker adaptation using Kullback Leibler divergence-based HMM which estimated the posterior probabilities of each deep neural network-based acoustic model. The system performance degraded due to the instability of developing lexical models. In [21], the authors discussed embedded-based recognition systems using empirical mode decomposition and the genetic algorithm. It reduced the computational load of the acoustic speech from intrinsic mode functions and further its estimated weight using a genetic algorithm to obtain the optimal solution. It does not apply to real-time systems.

The authors in [22] dealt with unknown unknowns from multistream speech recognition. Feedback-based techniques are utilized for deriving out-of-vocabulary. By doing so, the behavior of each signal is analyzed, and then, relevant information is obtained. Then, the Gaussian mixture model was used as a classifier, and thus, the

speech signal was recognized. This system is mostly dominated by probability errors. The study was enhanced by Zhang et al. [23] using a biologically inspired learning system. Global communication in the neural network estimated the synaptic weight of its corresponding presynaptic and postsynaptic neurons. It reduced the overhead of the hardware implementation with the least recognition rate. In [24], the authors studied a joint approach model for single-channel speaker identification systems. This model supported the fully blind system followed by a minimum mean square error of sinusoidal parameters. Code vector formulation from mixture estimation during the reconstruction stage degraded the signal quality.

In [25], the authors studied noise reduction by maximizing phoneme with enhanced accuracy. Though the system simplified the phoneme models by dropping the context dependence but the solutions become non-feasible. In [26], the authors discussed multilingual vision and the speech interaction process. The system estimated the eye gaze measure, head pose estimation, facial expression, and text-to-speech components and then fed into binary patterns of three orthogonal planes of shape domain. It was experimented in art critic and displayed better results. The time taken by the number of feature points is higher. In [27], the authors presented a gating neural networks model for large vocabulary audio–visual processes. The assessment of visual features affects the performance of the system, and it is resolved by introducing the gating layer [28] in neural networks. The system supports only 25% of features for systems training. The integration of audio and visual features remains a challenging task during performance analysis. The fusion models are developed for the feature fusion model, the decision fusion model, the multistream hidden Markov model [29] (HMM), the coupled HMM, and the turbo decoders. Lattice generation at each time frame creates a constrained maximization problem. In [30], the authors studied recurrent neural networks for multi-genre broadcast speech models by linear hidden network adaptation layer and the k-component adaptive layers. Feature representation models helped to achieve better speech classification with degraded speech perplexity and word error rate.

3 Research Methodology

This section presents the proposed model of the research study. The survey states that the analysis is done in automatic speech recognition (ASR) is still a challenging and demanding task. Due to variation in word utterance, this recognition system is in the upcoming research area. Feature extraction plays a vital role in ASR systems because the speech quality purely depends on it. Though variant models were suggested by the scholars, the accurate translation of the speech text from the speech signal is not yet achieved. Most of the analysis is done using hidden Markov models which throws issues like recognition error rate, accuracy, and speech perplexity. Here, we attempt to resolve the abovementioned issues and achieve better translation and recognition rates than others. The proposed phases are explained as follows:

3.1 Data Collection

This is the first step of the study. The news report dataset that comprises of audio and text records were collected from a public repository. Since it is a news dataset, the acquisition of relevant and irrelevant data is of higher, i.e., missing fields, irrelevant data entered irrelevant attributes.

3.2 Preprocessing

Preprocessing is one of the significant steps in our study. Here, a simple normalization technique is applied to both audio and text records. It is a process that alters the range of pixel intensity values. It is also known as contrast stretching or histogram stretching. The consistency of the signals is maintained by maximum and minimum pixel values under dynamic ranges. According to Eq. (1), the normalized values are computed as follows:

$$\text{In} = (I - \text{Min})\,[(\text{newMax} - \text{newMin})/\text{Max} - \text{Min}] + \text{newMin} \tag{1}$$

3.3 Feature Selection

This is the most and core part of the automatic speech recognition systems because it depicts the quality of the speech signal. Mel-frequency Cepstral Coefficient (MFCC) is employed for extracting the features from signals. Figure 1 presents the workflow of MFCCs.

Framing and blocking: The collected 1D signal is segmented into n frames of N samples. Likewise, the M samples in adjacent frames are overlapped by N–M samples. If the frame size is smaller than the sample sizes, then enough details are not acquired from the frame. Until the signal ends, the frames are processed.

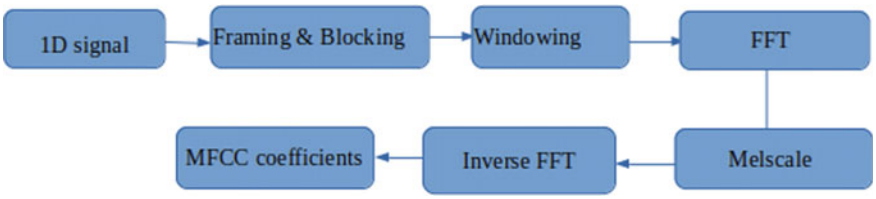


Fig. 1 Workflow of MFCCs

Windowing: In this step, the colliding of frames is minimized. Mostly, hamming windows are used for representing the input and output signal. Equation (2) presents the estimation of frames in windows.

$$\mathrm{Wn}(m) = 0.54 - 0.46\cos\left(2\pi m/\mathrm{Nm}^{-1}\right) \tag{2}$$

Fast Fourier Transform (FFT): It converts the spatial domain into the frequency domain. Each frame holds samples which convert into a frequency domain. It is one of the fastest algorithms which applies discrete Fourier transform (DFT).

$$D_k = \sum_{m=0}^{N_{m-1}} D_m e^{\frac{-j2km}{Nm}}, \quad \text{Where } k = 0, 1, 2 \ldots, \mathrm{Nm}^{-1} \tag{3}$$

Each DFT is computed separately for an easier computation process.

Mel Scale: Here, a triangular filter bank is utilized for computing energy functions. It is a set of bandpass filters that are being decided by steady Mel-frequency time. When frequency gets higher, the Mel space becomes wider. Mel-scaling mapping is done between observed frequency scales (Hz) and the perceived frequency scale (Mels). Finally, the first 13 coefficients were used, and the higher values were eliminated. These computed coefficients are then fed into DCT.

Discrete Cosine Transform (DCT): It helps to convert the log Mel spectrum into the spatial domain. The DFT is desirable for all coefficient estimation from DCT. The output obtained after applying DCT is known as Mel-frequency Cepstral Coefficient.

$$C_n = \sum k\, k - 1(\log\ Dk)\cos[m(k-1)\pi \tag{4}$$

where $m = 0, 1 \ldots k - 1$, C_n = MFCC value, m = no. of coefficients. The MFCC converts the input signal from 2 to 1D signal. These 1D signals are categorized into frames, and the neighboring frames are represented as ($M < N_m$). By doing so, the loss of pixel information is avoided. At windowing, the frequency domain is achieved by FFT. Then, the estimation of the magnitude spectrum is further transformed by Mel-frequency. For audio signals, the features retrieved are coefficients, delta, and the delta–delta. Likewise, the text features are mean, entropy, variance, skewness, kurtosis, and the root mean square.

3.4 Classification

It is one of the core parts which takes features, as input, to CNN, and then, speech is recognized in text form. It comprises three layers, namely the input layer, an output

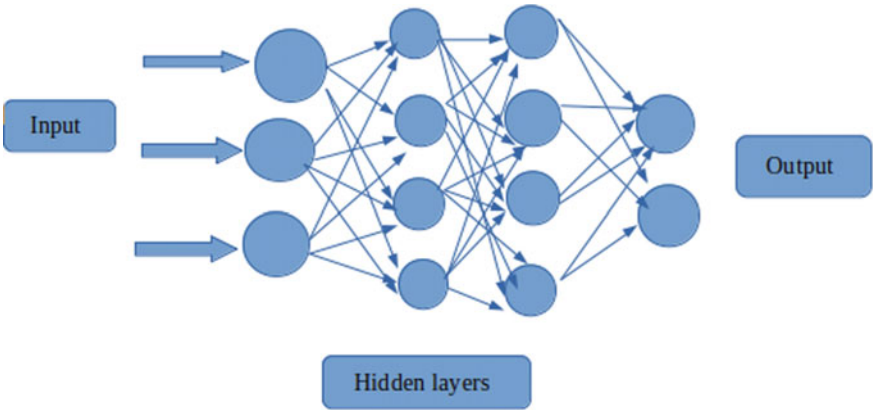


Fig. 2 Working of CNN

layer, and several hidden layers. Each layer has its functioning process with learning features. The most common layers in CNN are convolution, activation or ReLU, and pooling.

Convolution: A set of convolutional filters are used for activating certain features.

Rectified Linear Unit (ReLU): It maps the negative values to zero and administers the positive values. These are further processed onto activation layers.

Pooling layer: It identifies the variant features by nonlinear downsampling (Fig. 2).

Once after receiving the learning features, the CNN helps for classification. Since it is a connected layer, the k-dimensions define the number of classes. Figure 3 represents the workflow of our study.

4 Experimental Results

This section presents the experimental analysis of the proposed model. The proposed CNN classifier is experimented using Matlab, programming language. Initially, the data are being collected from a public repository. The confusion matrix is generated for recognizing its relevant classes. The performance of the CPU is taken while training the features. Table 1 presents the sample values of the performance of the CPU. It depicts that the minimized loss rate helps in achieving a better accuracy rate.

Figure 4 presents the confusion matrix between existing ANN and proposed CNN. The matrix analysis depicts that loss incurred by CNN is better than ANN. Region of convergence curve (ROC) defines the area covered by the bounded region. And Fig. 5 presents the comparison of ROC between proposed CNN and existing ANN.

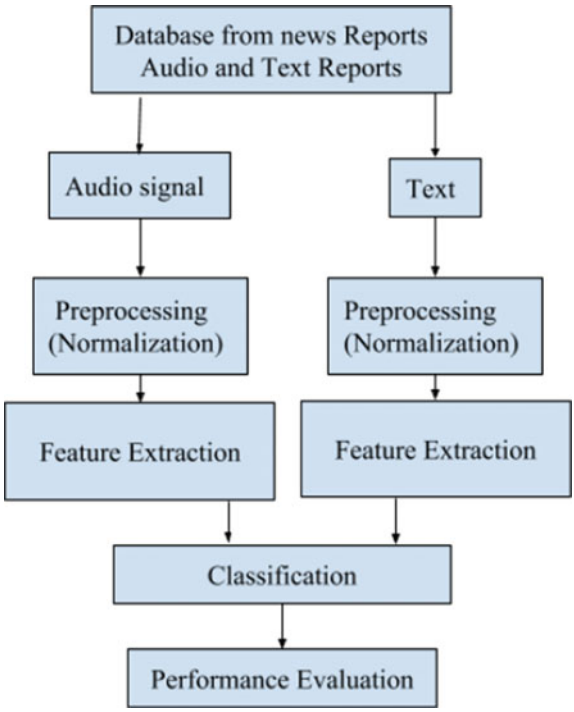


Fig. 3 Proposed workflow

Table 1 Performance of CPU for training the features

Epochs	Iteration	Elapsed time (s)	Accuracy (%)	Loss	Learning rate
1	1	0	0	1.6094	0.0100
1	50	0	100	0.2267	0. 0100
1	100	1	0	3.3115	0.0100
1	150	1	100	0. 1779	0.0100
1	200	2	0	2.6184	0.0100

It depicts that proposed CNN intelligently classifies its class better than the existing ANN. The performance metrics studied are explained as follows:

4.1 F-measure

It conveys the balance between precision and recall which is given in Eq. (5).

$$F-\text{measure}: 2 * ((\text{precision} * \text{recall})/(\text{precision} + \text{recall})) \tag{5}$$

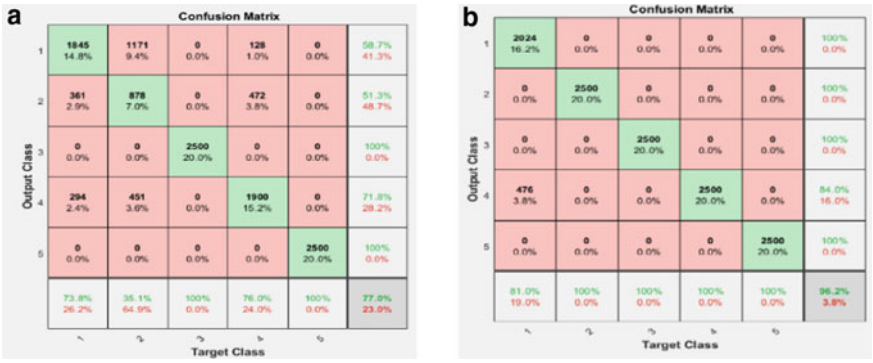


Fig. 4 Confusion matrix between (a) existing ANN and (b) proposed CNN

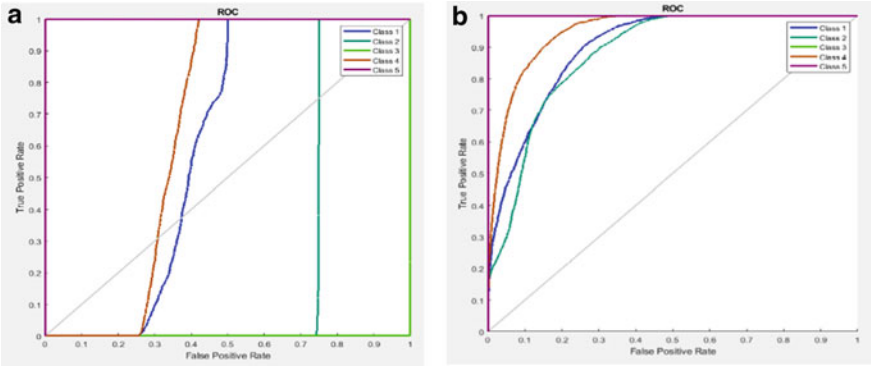


Fig. 5 Region of Convergence (a) existing ANN, (b) Proposed CNN

4.2 Accuracy

It defines the ability to distinguish normal and abnormal cases. It is given as:

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{6}$$

4.3 Sensitivity

It defines the ability to distinguish abnormal samples correctly. It is given as:

$$\text{Sensitivity} = (\text{TP})/(\text{TP} + \text{FN}) \tag{7}$$

Table 2 Performance comparison between existing ANN and proposed CNN

Classification	Accuracy	Sensitivity	Specificity
CNN	96.170	80.84	100
ANN	76.98	73.80	77.78

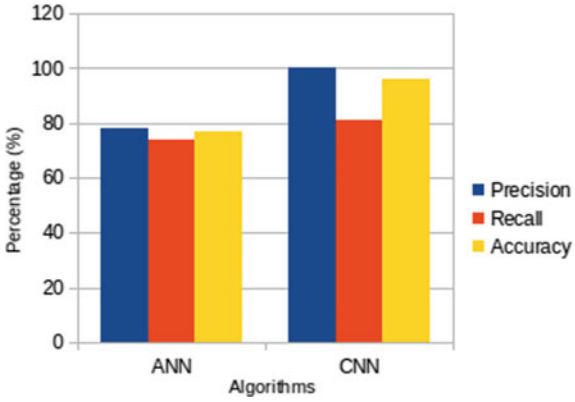


Fig. 6 Performance graph between existing ANN and proposed CNN

4.4 Specificity

It defines the ability to estimate normal samples correctly. It is given as (Table 2; Fig. 6):

$$\text{Specificity} : (\text{TN})/(\text{TN} + \text{FP}) \tag{8}$$

5 Conclusion

The recent innovation in information processing systems combined with intelligent models has gained much interest among the researchers. As we know, speech is a complex process that demands the coordination of articulation, breathing, and facial expression. This paper attempts to intelligently recognize the text from speech signals using improvised convolutional neural networks. The required data are collected from a public repository, news report dataset which generally comprises irrelevant data. It is being preprocessed by min–max normalization techniques that efficiently normalize the pixel intensity values of the data. The normalized data are then applied using MFCC that depicts the relevant features of audio and text data. The features extracted from audio are coefficients, delta, and delta-delta, and the text features are mean, entropy, variance, skewness, kurtosis, and the root mean square. These extracted features are fed into convolutional neural networks (CNNs), and the relevant signals

are classified. The proposed model is investigated over the news report dataset that comprises of two sources of data, audio, and text. The proposed CNN is compared with ANN which proves the efficiency in terms of accuracy, sensitivity, specificity, precision, recall, f-measure, and the Gaussian mean (Gmean). The accuracy of the proposed CNN 96% achieved is better than ANN, 76%.

References

1. Mikolov T, Karafiát M, Burget L, Cernocký J, Khudanpur S (2010) Recurrent neural network-based language model. In: Proceedings of the 11th annual conference of the international speech communication association, 2010, pp 1045–1048
2. Mikolov T, Zweig G (2012) Context-dependent recurrent neural network language model. In: Proceedings of the IEEE workshop spoken language technology, 2012, pp 234–239
3. Alumäe T (2013) Multi-domain neural network language model. In: Proceedings of the 14th annual conference of the international speech communication association, 2013, pp 2182–2186
4. Chen X, Wang Y, Liu X, Gales MJ, Woodland PC (2014) Efficient GPU-based training of recurrent neural network language models using spliced sentence bunch. In: Proceedings of the 11th annual conference of the international speech communication association, 2014, pp 641–645
5. Chen X et al (2015) Recurrent neural network language model adaptation for multi-genre broadcast speech recognition. In: Proceedings of the 6th annual conference of the international speech communication association, 2015, pp 3511–3515
6. Hawley MS, Cunningham SP, Green PD, Enderby P, Palmer R, Sehgal S, O'Neill P (2013) A voice-input voice-output communication aid for people with severe speech impairment. IEEE Trans Neural Syst Rehabil Eng 21(1):23–31
7. Shao Y, Chang CH (2011) Bayesian separation with sparsity promotion in perceptual wavelet domain for speech enhancement and hybrid speech recognition. IEEE Trans Syst Man Cybern 41(2)
8. Meltzner GS, Heaton JT, Deng Y, De Luca G, Roy SH, Kline JC (2017) Silent speech recognition as an alternative communication device for persons with laryngectomy. IEEE/ACM Trans Audio Speech Lang Process 25(12):2386–2398
9. Wu B, Li K, Ge F, Huang Z, Yang M, Siniscalchi SM (2017) An end-to-end deep learning approach to simultaneous speech dereverberation and acoustic modelling for robust speech recognition. IEEE J Sel Top Sign Process 11(8):1–11
10. Baby D, Virtanen T, Gemmeke JF, Van Hamme H (2015) Coupled dictionaries for exemplar-based speech enhancement and automatic speech recognition. IEEE/ACM Trans Audio Speech Lang Process 23(11):1788–1799
11. Chandrakala S, Rajeswari N (2016) Representation learning-based speech assistive system for persons with dysarthria. IEEE Trans Neural Syst Rehabil Eng 1–12
12. Shahamiri SR, Salim SSB (2014) A multi-views multi-learners approach towards dysarthric speech recognition using multi-nets artificial neural networks. EEE Trans Neural Syst Rehabil Eng 22(5):1053–1063
13. Mitra V, Nam H, Espy-Wilson CY, Saltzman E, Goldstein L (2011) Articulatory information for noise robust speech recognition. IEEE Trans Audio Speech Lang Process 19(7):1913–1924
14. Sahraeian R, Van Compernolle D (2017) Cross-lingual and multilingual speech recognition based on the speech manifold. IEEE/ACM IEEE Trans Audio Speech Lang Process 25(12):2301–2312
15. Grozdić ĐT, Jovičić ST (2017) Whispered speech recognition using deep denoising autoencoder and inverse filtering. IEEE/ACM Trans Audio Speech Lang Process 25(12):2313–2322

16. Ming J, Crookes D (2017) Speech enhancement based on full-sentence correlation and clean speech recognition. IEEE/ACM Trans Audio Speech Lang Process 25(12):1–13
17. Lee J, Park S, Hong I, Yoo H-J (2016) An energy-efficient speech extraction processor for robust user speech recognition in mobile head-mounted display systems. IEEE Trans Circ Syst 64(4):1–13
18. Siniscalchi SM, Yu D, Deng L, Lee CH (2013) Speech recognition using long-span temporal patterns in a deep network model. IEEE Sign Process Lett 20(3):1–12
19. Bapat OA, Fastow RM, Olson J (2013) Acoustic coprocessor for HMM-based embedded speech recognition systems. IEEE Trans Consum Electron 59(3):629–633
20. Kim M, Kim Y, Yoo J, Wang J (2017) Regularized speaker adaptation of KL-HMM for dysarthric speech recognition. IEEE Trans Neural Syst Rehabil Eng 25(9):1–14
21. Pan S-T, Li X-Y (2012) An FPGA-based embedded robust speech recognition system designed by combining empirical mode decomposition and a genetic algorithm. IEEE Trans Instrum Meas 61(9):2560–2572
22. Hermansky H (2013) Multistream recognition of speech: dealing with unknown unknowns. IEEE Proc 101(5):1076–1088
23. Zhang Y, Li P, Jin Y, Choe Y (2015) A digital liquid state machine with biologically inspired learning and its application to speech recognition. IEEE Trans Neural Netw Learn Syst 26(11):2635–2649
24. Mowlaee P, Saeidi R, Christensen MG, Tan Z-H (2012) A joint approach for single-channel speaker identification and speech separation. IEEE Trans Audio Speech Lang Process 20(9):2586–2601
25. Petkov PN, Henter GE, Bastiaan Kleijn W (2013) Maximizing phoneme recognition accuracy for enhanced speech intelligibility in noise. IEEE Trans Audio Speech Lang Process 21(5):1035–1045
26. Reale MJ, Liu P, Yin L (2013) Art critic: multi signal vision and speech interaction system in a gaming context. IEEE Trans Cybern 43(6):1546–1559
27. Tao F, Busso C (2017) Gating neural network for large vocabulary audiovisual speech recognition. IEEE Trans Audio Speech Lang Process 21(5):1–12
28. Deena S, Hasan M, Doulaty M, Saz O, Hain T (2019) Recurrent neural network language model adaptation for multi-genre broadcast speech recognition and alignment. IEEE/ACM Trans Audio Speech Lang Process 27(3):572–578
29. Abdelaziz AH (2018) Comparing fusion models for DNN-based audiovisual continuous speech recognition. IEEE Trans Audio Speech Lang Process 26(3):475–484
30. Yoon WJ, Park KS (2011) Building robust emotion recognition system on heterogeneous speech databases. IEEE Proc 57(2):747–750

An IoT-Based Temperature Measurement Platform for a Real-Time Environment Using LM35

Anupam Kumar Sharma, Prashant Singh, Dhyanendra Jain, Anjani Gupta, and Prashant Vats

Abstract The number given to an object to signify its warmth is called its temperature. People tried to quantify and measure differences in warmth, which led to the invention of the temperature notion. When a hot item collides with a cold item, heat is transferred until the two things are the same temperature. The two items are considered to be in thermal equilibrium once the heat transfer is complete. The quantity of warmth that is the same for two or more things in thermal equilibrium is thus defined as temperature. In this study, we are presenting a microcontroller system that will automatically estimate the temperatures of certain area or surroundings using the sensing devices LM35. The results of a rooms or atmosphere-related process is seen or considered.

Keywords LM35 · Sensors · IoT · Smart homes · Microcontrollers · Temperature sensing

1 Introduction

The number given to an object to signify its warmth is called its temperature. People tried to quantify and measure differences in warmth, which led to the invention of the temperature notion. When a hot item collides with a cold item, heat is transferred until the two things are the same temperature. The two items are considered to be in thermal equilibrium once the heat transfer is complete. The quantity of warmth that is the same for two or more things in thermal equilibrium is thus defined as temperature. The triple point of water is what we call it. The digital thermometer combines components such as a microcontroller and an LM35 temperature sensor, all of which are programmed in the embedded C programming language. Digital thermometers may now be used in home automation; IoT services for medical records, industrial activities, and many more applications thanks to technological advancements. Some facts

A. K. Sharma (✉) · P. Singh · D. Jain · A. Gupta · P. Vats
Dr.Akhilesh Das Gupta Institute of Technology and Management, Guru Gobind Singh Indraprastha University, New Delhi, India
e-mail: anupam.sharma@adgitmdelhi.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_57

were not in place at the development of this thermometer, resulting in its shortcomings, and a digital thermometer was created as a result of technological progression. The digital thermometer combines components such as a microcontroller and a LM35 temperature sensor, all of which are programmed in the embedded C programming language. Digital thermometers may now be used in home automation; IoT services for medical records, industrial activities, and many more applications thanks to technological advancements. The LM35 is a type of sensor with a precession voltage that changes with the room temperature. It is a tiny, cheap IC capable of measuring −55 to 150° Centigrade temperatures. It may be simply coupled to almost any controller with ADC functionality and to any development platform, such as Arduino. Attach the Chip port to the earth of the circuitry and give a Cc input line-controlled supply, like 5 V dc (VS). As shown in Fig. 1, the temperatures may now be measured using power output. When the temp is 0° Temperature, the power output would be 0 V. A surge of latent viral (10 mV) would occur with each grade Centigrade temperature rise. Using the formulas below, we can convert voltage to temperature.

$$V_{\text{out}} = (10\,\text{mv})/C\hat{}\text{o}*T \tag{1}$$

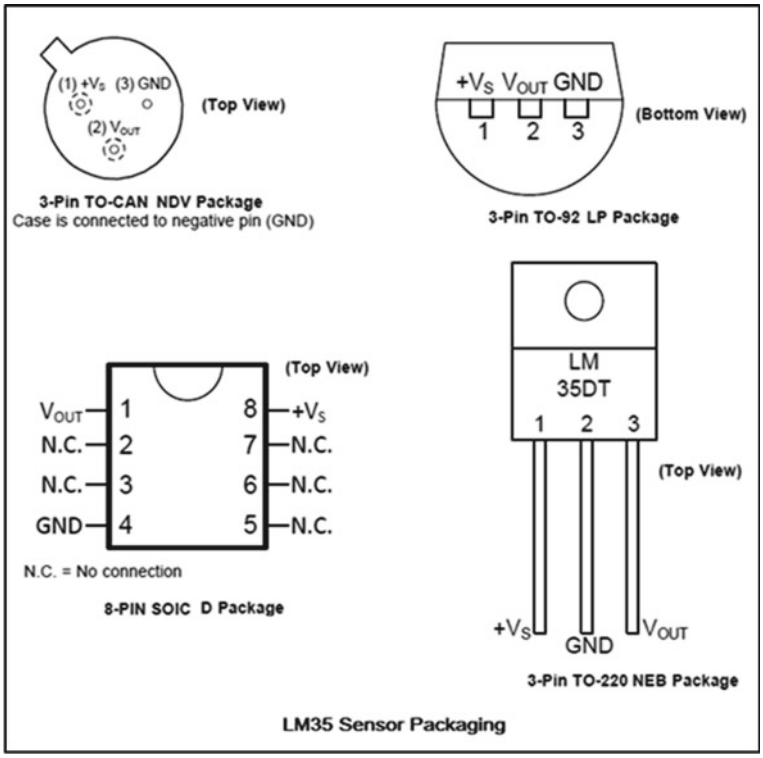


Fig. 1 To show the LM35 sensor IC packaging

The degree of output current is the LM35 output signal and T would be the temperature observed.

2 Related Work

This part reviews the literature on by providing a comparative survey of the related work on temperature controlling using LM35 sensor is given in Table1.

3 LM35 Microcontroller-Based Temperature Sensing

AC voltage in accordance to the present temperature is generated by the temperature probe LM35. A thermal measuring of degrees c may easily convert the electrical output. The advantage of LM35 would be that it does not have to be manually adjusted. Readings between −55 and 150 °C Temperature may be found in the LM35 based microcontroller. The precision level is exceptionally high when operated at the appropriate temperature and relative humidity conditions. The voltage production to Deg c. may simply and easily be adapted. In order to better understand the principles of operation of the LM35, the sequential scale factor must always be comprehended. It shows that the temperatures value is increased by one before the vdd pin sensor is 10 mv. Two separate circuits can be connected to the LM35. Both deliver separate results. The initial setup can only monitor affirmative temperature around 2 °C and 150 °C. As shown in Fig. 2, device LM35 is solely connected, and the output is directly linked to the analogue with the digital converter topologies in this first setup. As shown in Fig. 3. We could employ all of the sensing resources in the subsequent configuration and detect the full mean temperature from −55 to 150 °C.

4 Proposed Methodology

The warmth in our suggested work is monitored using an Arduino board and an LM35. The diagram of the working circuit is shown in Fig. 4, and the receiving pin is the Analogue port A0 from the Arduino Uno. Then, we are going to attach to it the microcontroller analogue signal pin. We employ a looping functionality to constantly monitor and afterwards publish a temperature around + 5 voltages on the microcontroller digital output pin and grounding the Gnd line on a 162-character LCD. The Centigrade temp is translated to Fahrenheit, following the technique indicated in Eq. (2) by entering once more value at 162 LCD in Fahrenheit. An Arduino has a 4-bit mode 162 pins LCD connection. The circuitry schematic of the suggested work is presented in Fig. 5.

Table 1 A comparative survey of the related work on temperature controlling using LM35 sensor

S. No	Authors	Key points	Merits	Demerits	Tools used	References
1	Barik (2019)	HTU 211D sensor device, Raspberry Pi, Arduino UNO, and ESP8266 Wi-Fi module are used in this system	Provides consumers with the ability to evaluate environmental changes and take relevant action + It is low-cost and uses a lot less power	It cannot be used to figure out how much moisture is in the soil or to monitor a farm	HTU 211D sensor device, Raspberry Pi, Arduino UNO, and ESP8266 Wi-Fi	[1]
2	Subhasini (2019)	To concentrate on how FPGA-based hardware resources IoT-based temperature monitoring accessible from anywhere	Communication protocols, data acquisition, and control systems are all included in the IoT-based FPGA	Not defined	FPGA starting kit Spartan3AN with LM35 temperature sensor	[2]
3	Akash (2017)	On a real-time basis, controlled a single channel relay, which may be utilised to manage water flow on the field	Allows the user to save time and effort by analysing accurate changes in the environment and taking appropriate action	Not defined	DHT11 sensor, Raspberry Pi modal 3 B with Internet connectivity	[3]
4	Utomo (2019)	Using a temp. and humidity sensor (DHT11), Raspberry Pi, Arduino module, to monitor temperature of a specific room	If temperature and humidity in server room surpass defined thresholds, system send telegram signals to adjust the air conditioner temperature	Not defined	Temp. and Humidity Sensor (DHT11), Raspberry Pi, Arduino module	[4]

(continued)

Table 1 (continued)

S. No	Authors	Key points	Merits	Demerits	Tools used	References
5	Lavanya (2016)	Temperature and humidity data are retrieved, and these data are felt and communicated to the Internet	With the automatic temperature and humidity management, the expense of monitoring the system is reduced	Not Defined	Raspberry Pi, a Wi-Fi dongle, a DHT11 sensor, and an Android phone	[5]
6	Chen (2019)	Detecting the activation of ventilation equipment in subterranean tunnels and utilising a gateway to receive sensed data to monitor tunnel temperature dispersion	Offers an IoT-based underground cable temperature monitoring system and tunnels for accurately and consistently detect temperature in the tunnel in real time and send the data to a gateway	Not defined	LM35 as a temperature sensor	[6]
7	Muhammad (2018)	By transferring patient information to clinicians, the suggested architecture of the patient management system may be utilised to monitor health-related concerns	Keep track of health issues by transferring patient information to doctors	Not defined	LM35 temperature sensor + Wi-Fi module ESP8266 + GSM module SIM900D	[7]

(continued)

$$(\text{Temperature C}) * 9.0/5.0) + 32.0 \quad (2)$$

Table 1 (continued)

S. No	Authors	Key points	Merits	Demerits	Tools used	References
8	Tamilselvi (2020)	The accelerometer sensor was utilised to depict the coma patients' bodily movement	Offers coma patient monitoring system based on IoT and GSM	Not defined	Eyeblink sensor and SPO2 sensor + Arduino Software	[8]
9	Karim (2018)	To decrease man monitoring and to build IoT-based real-time temperature and humidity monitoring system, DHT-11 and ESP-8266 Node MCU module are used	Controlling and monitoring the humidity and temperature influence on raw foods using IoT to improve food safety	Cost of the system increases with integrating it to larger areas	NodeMCU ESP8266 Wi-Fi + Arduino	[9]
10	Bharath Prabu (2020)	LabVIEW is used to keep track of and control temperature	Monitoring and adjusting temperature to run chillier in businesses using data from the server	Not defined	LM35 sensor + LabVIEW software	[10]

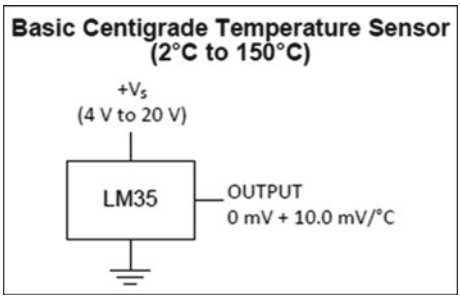


Fig. 2 To show the LM35 as basic centigrade temperature sensor

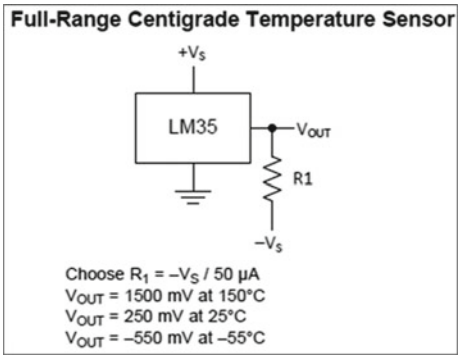


Fig. 3 To show the LM35 as full range centigrade temperature sensor

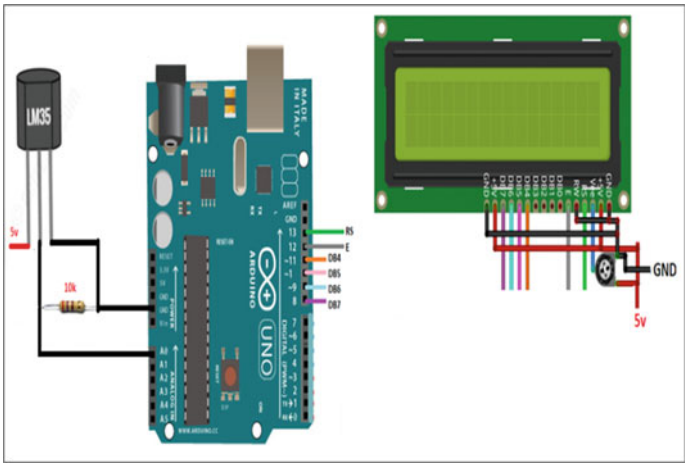


Fig. 4 To illustrate Arduino Uno—measuring warmth—the LM35 thermal sensor 16 to 2 liquid crystal display

5 Experimental Setup and Results

The LM35 is a temperature sensor that is both analogue and linear. The output voltage is therefore proportional to the temperature. For every 1 °C increase in temperature, the output voltage climbs by 10 mV. From 0 to 5 V, the Arduino can read it. This is saved as a ten-bit number by the Arduino (0–1023) as shown in Fig. 6. Figure 7 is used to show the connection made for LM35. Figure 8. Feeding the code to Arduino Uno. Figure 9 change in the temperature value. Examine the temperature with the serial monitor open. Every 0.5 s, the temperature will fluctuate.

```
#include<LiquidCrystal.h>
LiquidCrystal lcd(12,11,10,9,8,7);
const int kPinTemp = A0;
 void loop()
 {
 float reading = analogRead(kPinTemp);   //Analog pin reading output
voltage by Lm35
 float temperatureC=kPinTemp/2.048;      //Finding the true
centigrade/Celsius temperature
 lcd.setCursor(0,0);                 //Place lcd cursor at coulomb-1 row-1
 lcd.print("CENTI TEMP= ");
 lcd.print(temperatureC);              //Print centigrade temperature on lcd
 float temperatureF= ((temperatureC ) * 9.0 / 5.0) + 32.0;
 lcd.setCursor(0,1);                 //Place lcd cursor at coulomb-2 row-2
 lcd.print("FARE TEMP= ");
 lcd.print(temperatureF);              //Print Fahrenheit temperature on lcd
 delay(800);                    //Delay to remain data appear on lcd.
 lcd.clear();                   //Clear lcd screen
 }
```

Fig. 5 Code for measuring-room—temperature-with LM35—using Arduino Uno

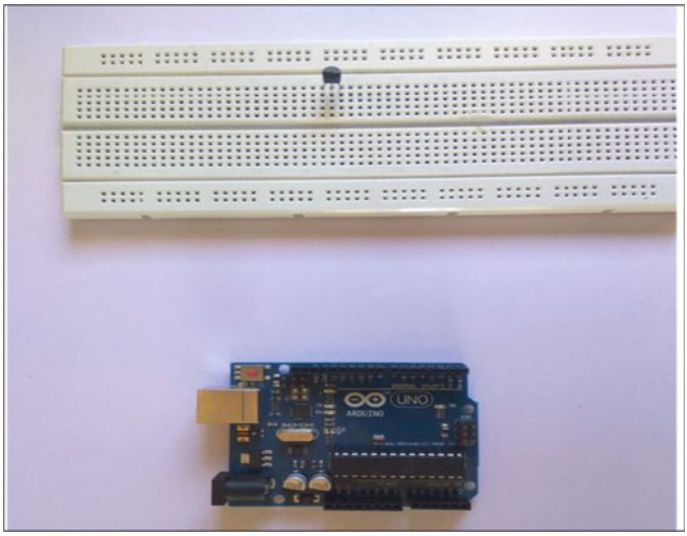


Fig. 6 To show LM35 is a temperature sensor

6 Conclusions

This study developed an Arduino system, which can autonomously detect the warmth of a certain surroundings or area utilising the Senior LM35s. The result of a space or atmosphere-related process is seen or considered. The temperatures in our suggested

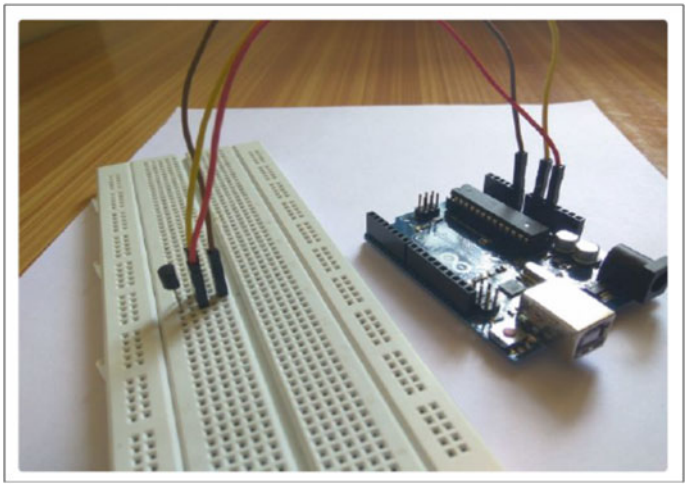


Fig. 7 To show the connection made for LM35

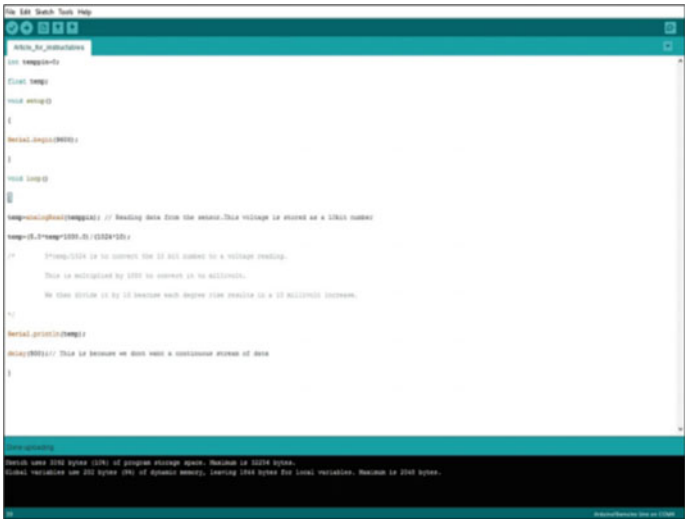


Fig. 8 Feeding the code to Arduino Uno

work were monitored using an Arduino board and an LM35. A measuring temperature in C can be read easily for the output voltage. The benefit of LM35 over thermocouple is that no additional adjustment is required.

Fig. 9 The change in the temperature value

References

1. Barik L (2019) IoT based temperature and humidity controlling using arduino and raspberry Pi. Int J Adv Comput Sci Appl 10(9) (2019)
2. Subashini G (2019) IOT based temperature monitoring system using FPGA. J Analog Digital Devices 4(3):2019
3. Akash AB (2017) IoT-based temperature and humidity monitoring system for agriculture. Int J Innov Res Sci Eng Technol 6(7) (2017)
4. Utomo MA (2019) Server room temperature and humidity monitoring based on Internet of Thing (IoT). In: IOP conference series: Journal of Physics vol 1306
5. Lavanya M (2016) Iot based automated temperature and humidity monitoring and control. J Chem Pharm Sci (5)
6. Chen MF (2019) An IoT-based temperature monitoring system for underground cable tunnels. In: Proceedings of the 2019 IEEE PES GTD Asia, 2019
7. Muhammad S (2018) IOT Based health monitoring system using Arduino UNO, GSM SIM900D module and Wi-Fi module ESP8266. Int J Sci Eng Res 9(12)
8. Tamilselvi V (2020) IOT based health monitoring system. In: 6th IEEE international conference on advanced computing and communication systems (ICACCS), 2020
9. Karim AB (2018) Monitoring food storage humidity and temperature data using IoT. MOJ Food Process Technol 6(4):400–404
10. Bharath Prabu R (2020) IoT based temperature monitoring and controlling system by using my RIO-LabVIEW. Int J Innov Res Electr Electron Instrum Control Eng 8(4)

Cross Validated Canonical Variate Analysis for Damage Grading After Math of Earthquake Building

M. Shyamala Devi, A. Peter Soosai Anandarai, M. Himavanth Sai, S. Suraj, S. Saranraj, and M. Guruprasath

Abstract Seismic events are disasterous happenings that can destroy structures and endanger human life. Due to the obvious innumerable sources of environmental uncertainty, estimating earthquake damage to buildings remains a difficult task. A variety of factors influence the severity of earthquake-induced building damage, such as amplitude, proximity to geographic centre, geological features and compliance with building reliability. Examining the damage rate of concrete structural system is critical for acknowledging the maintenance need avoiding the level of damage during the next catastrophic event. With this incentive, this paper discusses the significance of building damage characteristics in predicting building damage grade. The building earthquake damage dataset with 27 building attributes with 762,106 building damage details from KAGGLE warehouse is used in the execution analysis to determine the grade of damage. The earthquake damage dataset is preprocessed with missing values estimation, categorical feature encoding and feature scaling. The dataset with all 26 features is applied to all the classifiers to grade the damage with and without components scaling and the performance is analysed cross validating the training and testing data with 80:20, 70:30 and 60:40. The dataset is minimized with 10 linear discriminant components and then applied to all the classifiers to grade the damage with and without components scaling and the performance is analysed with cross validating the training and testing data with 80:20, 70:30 and 60:40. The scripting is written in Python and implemented with Anaconda Navigator, and the outcome shows that the random forest classifier is exhibiting 92% of accuracy after feature scaling and giving 96% of accuracy with 10 component LDA reduced dataset before and after feature scaling.

Keywords Machine learning · Classifier · Scaling · Encoding · Cross validation

M. Shyamala Devi (✉) · A. Peter Soosai Anandarai · M. Himavanth Sai · S. Suraj · S. Saranraj · M. Guruprasath
Computer Science and Engineering, Vel Tech Rangarajan Dr, Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India
e-mail: shyamaladevim@veltech.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_58

1 Introduction

Earthquakes are cataclysmic events that can destroy buildings and threaten human well-being. The severity of quake-activated structure damage is determined by a variety of factors such as size, distance from the focal point and land conditions, as well as seismic structure execution. The ability to quickly assess the distribution pattern and severity of building damage is essential for post event emergency response and recovery. Building damage grade categorization takes a substantial period of time and personnel resources, and it can take several months after the earthquake. The time of occurrence of a potential tremor is erratic. As a result, seismic risk assessment is critical for preparing for appropriate tremor relief exercises. An exhibition-based quake designing evaluation technique has been developed to provide a seismic dynamic key arrangement. A comprehensive vibrational restructuring strategy is a complex computation that explains how to enhance a building's seismic presentation.

2 Literature Review

This paper [1] explored the fictional scenario weakness bends in top ground velocity increases and seismographic power using the numerical model of building harm rate by tremors taking into account both the quake blockage limit and the seismic activity power. This study [2] clarified the important technological issues associated with post-earthquake emergency building inspection procedures are presented in this paper, and recommendations are given based on a comprehensive specific experience from previous earthquakes. This study [3] performs pre and post event, Normalized Digital Surface Models, examination presentation was assessed by a precision evaluation technique utilizing reference ground truth raster data.

This study [4] investigated how image preparation and PC vision advancements were used for seismic tremor harm recognition. For a delegated harm location measure, a city model was used. Visual damage exploration of building structures was completed by looking at pre and post event images in metropolitan regions. The prototype [5] is an innovative rough set-based artificial neural network strategy that is used to predict the vibration information levels captured by the sensors in the damaged building. This paper [6] reviewed the construction hazardous and to recognize the constructions influenced that identify a hazard for the neighbourhood, as well as the protected advancements that may be involved or utilized as transient sanitariums following a moderate and severe quake.

This research evaluated the security problems present both in the event of seismic activity and in primary response, quake danger expectation must be treated probabilistically. Two quantifiable techniques are presented in this review for the prediction of likely earthquake—induced harm to low and mid-ascent built-up substantial structures [7]. This paper proposes a novel convolutional neural networks [8] method in

conjunction with ordinal recurrence targeting for assessing the level of structure damage caused by seismic events using aerodynamic metaphors. This study [9] posited a further parchment device that addresses the need for a global but itemized procedure for building damage analysis post-seismic aftershocks. The multi threat system [10] evaluates the injury hazard of an elevated surface exposed to uncontrollable shaking and wind risks separately and instantaneously, and is divided into three sections: danger display, underlying delicacy examination, and harm likelihood calculation.

3 Our Contributions

The overall architecture of the work is shown in Fig. 1. The following contributions are provided in this work.

- Firstly, the earthquake damage dataset is preprocessed with missing values estimation, categorical feature encoding and feature scaling.
- The dataset with all 26 features is applied to all the classifiers to grade the damage with and without components scaling and the performance is analysed cross validating the training and testing data with 80:20, 70:30 and 60:40.
- The dataset is minimized with 10 and 15 linear discriminant components and then applied to all the classifiers to grade the damage with and without components scaling and the performance is analysed with cross validating the training and testing data with 80:20, 70:30 and 60:40. The metrics used for analysis are precision, recall, F-score, accuracy and execution time.

4 Implementation Setup

The earthquake damage dataset with 762,106 rows and 26 feature attributes from KAGGLE repository is subjected with the data preprocessing such as refilling the missing values, encoding of the categorical variables. The dataset information is shown in the Fig. 2.

4.1 Prescriptive and Predictive Damage Grade Analysis

The distribution of the target variable damage grade is shown below in Fig. 3

The predictive analysis of the dataset with entire 26 features is applied to all the classifiers to grade the damage with and without feature scaling and the performance is analysed with cross validating the training and testing data with 80:20, 70:30 and 60:40 and is shown in Tables 1, 2 and 3 and Fig. 4. Table 1 gives a summary of all heading levels.

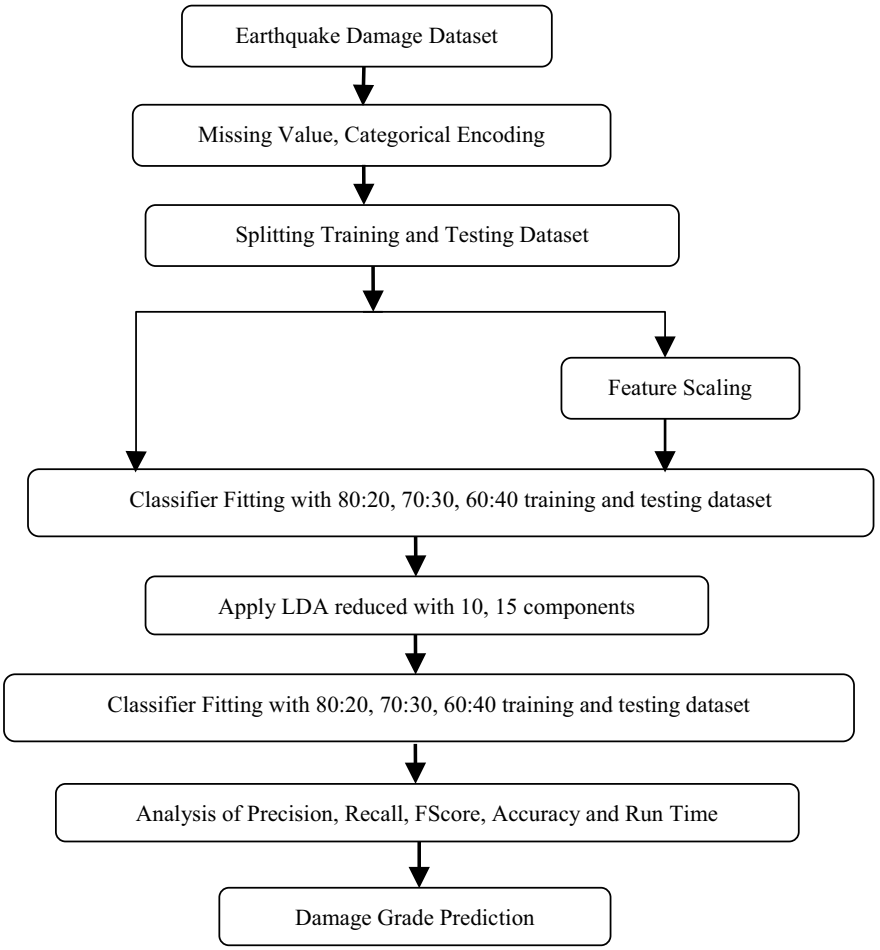


Fig. 1 Architecture system workflow

The dataset is applied with linear discriminant analysis and is reduced with 10 component features. The 10 component LDA reduced dataset is applied to all the classifiers to grade the damage with and without feature scaling and the performance is analysed with cross validating the training and testing data with 80:20, 70:30 and 60:40 and is shown in Tables 4, 5 and 6 and Fig. 5.

5 Mathematical Modelling and Analysis

The earthquake damage dataset with 26 independent features and 1 dependent feature "damage grade" is represented as in Eq. (1),

```
 0   count_floors_pre_eq                     762106 non-null  int64
 1   count_floors_post_eq                    762106 non-null  int64
 2   age_building                            762106 non-null  int64
 3   plinth_area_sq_ft                       762106 non-null  int64
 4   height_ft_pre_eq                        762106 non-null  int64
 5   height_ft_post_eq                       762106 non-null  int64
 6   land_surface_condition                  762106 non-null  int64
 7   foundation_type                         762106 non-null  object
 8   roof_type                               762106 non-null  int64
 9   ground_floor_type                       762106 non-null  int64
 10  3_floor_type                            762106 non-null  int64
 11  position                                762105 non-null  float64
 12  plan_configuration                      762105 non-null  object
 13  has_superstructure_adobe_1              762106 non-null  int64
 14  has_superstructure_1_mortar_stone       762106 non-null  int64
 15  has_superstructure_stone_flag           762106 non-null  int64
 16  has_superstructure_cement_mortar_stone  762106 non-null  int64
 17  has_superstructure_1_mortar_brick       762106 non-null  int64
 18  has_superstructure_cement_mortar_brick  762106 non-null  int64
 19  has_superstructure_5                    762106 non-null  int64
 20  has_superstructure_bamboo               762106 non-null  int64
 21  has_superstructure_4_non_engineered     762106 non-null  int64
 22  has_superstructure_4_engineered         762106 non-null  int64
 23  has_superstructure_3                    762106 non-null  int64
 24  condition_post_eq                       762106 non-null  object
 25  damage_grade                            762094 non-null  object
 26  technical_solution_proposed             762094 non-null  object
dtypes: float64(1), int64(21), object(5)
```

Fig. 2 Dataset information with its attributes

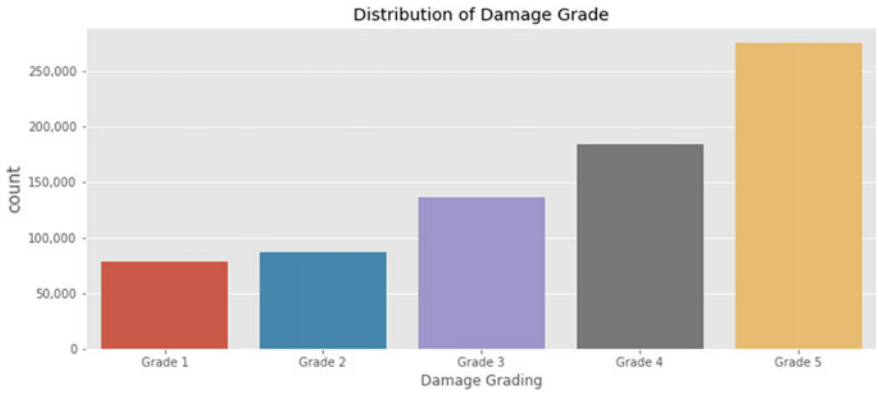


Fig. 3 Damage grade target distribution

Table 1 Classifier performance indices with 80:20 before and after feature scaling

Classifier	Before feature scaling					After feature scaling				
	Precision	Recall	F-Score	Accu	Run Time	Precision	Recall	F-Score	Accu	Run Time
LReg	0.57	0.58	0.55	0.58	0.21	0.71	0.70	0.69	0.70	0.19
KNN	0.42	0.41	0.40	0.41	0.07	0.77	0.76	0.76	0.76	0.18
KSVM	0.31	0.34	0.28	0.34	1.26	0.83	0.79	0.78	0.79	0.60
GNB	0.58	0.30	0.29	0.30	0.00	0.58	0.25	0.23	0.25	0.00
Dtree	0.80	0.80	0.80	0.80	0.02	0.80	0.80	0.80	0.80	0.01
Etree	0.76	0.76	0.76	0.76	0.01	0.76	0.76	0.76	0.76	0.00
RFor	0.83	0.83	0.83	0.83	0.10	0.93	0.92	0.93	0.92	0.10
Ridge	0.68	0.65	0.63	0.65	0.08	0.68	0.65	0.63	0.65	0.00
RCV	0.68	0.65	0.63	0.65	0.04	0.68	0.65	0.63	0.65	0.02
SGD	0.78	0.43	0.36	0.43	0.13	0.67	0.67	0.66	0.67	0.10
PAg	0.42	0.45	0.37	0.45	0.07	0.55	0.51	0.52	0.51	0.03
Bagg	0.85	0.85	0.85	0.85	0.11	0.85	0.85	0.85	0.85	0.13

Table 2 Classifier performance indices with 70:30 before and after feature scaling

Classifier	Before feature scaling					After feature scaling				
	Precision	Recall	F-Score	Accu	Run Time	Precision	Recall	F-Score	Accu	Run Time
LReg	0.63	0.55	0.52	0.55	0.18	0.71	0.70	0.69	0.70	0.19
KNN	0.42	0.40	0.39	0.40	0.08	0.77	0.76	0.76	0.76	0.16
KSVM	0.30	0.32	0.26	0.32	1.02	0.83	0.79	0.78	0.79	0.58
GNB	0.60	0.31	0.30	0.31	0.01	0.58	0.25	0.23	0.25	0.00
Dtree	0.81	0.81	0.81	0.81	0.02	0.80	0.80	0.80	0.80	0.01
Etree	0.79	0.78	0.79	0.78	0.01	0.76	0.76	0.76	0.76	0.00
RFor	0.84	0.84	0.84	0.84	0.09	0.93	0.92	0.93	0.92	0.10
Ridge	0.68	0.65	0.63	0.65	0.01	0.68	0.65	0.63	0.65	0.01
RCV	0.68	0.65	0.63	0.65	0.01	0.68	0.65	0.63	0.65	0.01
SGD	0.59	0.27	0.14	0.27	0.14	0.68	0.67	0.65	0.67	0.07
PAg	0.55	0.19	0.13	0.19	0.04	0.48	0.47	0.45	0.47	0.01
Bagg	0.83	0.83	0.83	0.83	0.09	0.84	0.83	0.83	0.83	0.12

$$X = \{[x_1, x_2, x_3, \ldots, x_{26}], [Y]\} \tag{1}$$

The X represents the dataset, Y represents the target variable. The earthquake damage dataset is pre-processed with categorical feature encoding based on the distinct values in the respective attributes and is denoted in Eq. (2),

Table 3 Classifier performance indices with 60:40 before and after feature scaling

Classifier	Before feature scaling					After feature scaling				
	Precision	Recall	F-Score	Accu	Run Time	Precision	Recall	F-Score	Accu	Run Time
LReg	0.52	0.56	0.53	0.56	0.18	0.70	0.70	0.69	0.70	0.14
KNN	0.41	0.39	0.39	0.39	0.10	0.75	0.74	0.74	0.74	0.26
KSVM	0.31	0.31	0.24	0.31	0.80	0.81	0.78	0.77	0.78	0.42
GNB	0.57	0.29	0.28	0.29	0.01	0.57	0.24	0.21	0.24	0.00
Dtree	0.81	0.81	0.81	0.81	0.01	0.81	0.81	0.81	0.81	0.02
Etree	0.78	0.78	0.78	0.78	0.01	0.78	0.78	0.78	0.78	0.02
RFor	0.84	0.83	0.83	0.83	0.08	0.93	0.92	0.93	0.92	0.07
Ridge	0.68	0.66	0.64	0.66	0.01	0.68	0.66	0.64	0.66	0.00
RCV	0.68	0.66	0.64	0.66	0.03	0.68	0.66	0.64	0.66	0.02
SGD	0.57	0.34	0.22	0.34	0.12	0.64	0.64	0.63	0.64	0.07
PAg	0.64	0.30	0.16	0.30	0.04	0.60	0.60	0.59	0.60	0.03
Bagg	0.84	0.84	0.84	0.84	0.08	0.83	0.83	0.83	0.83	0.10

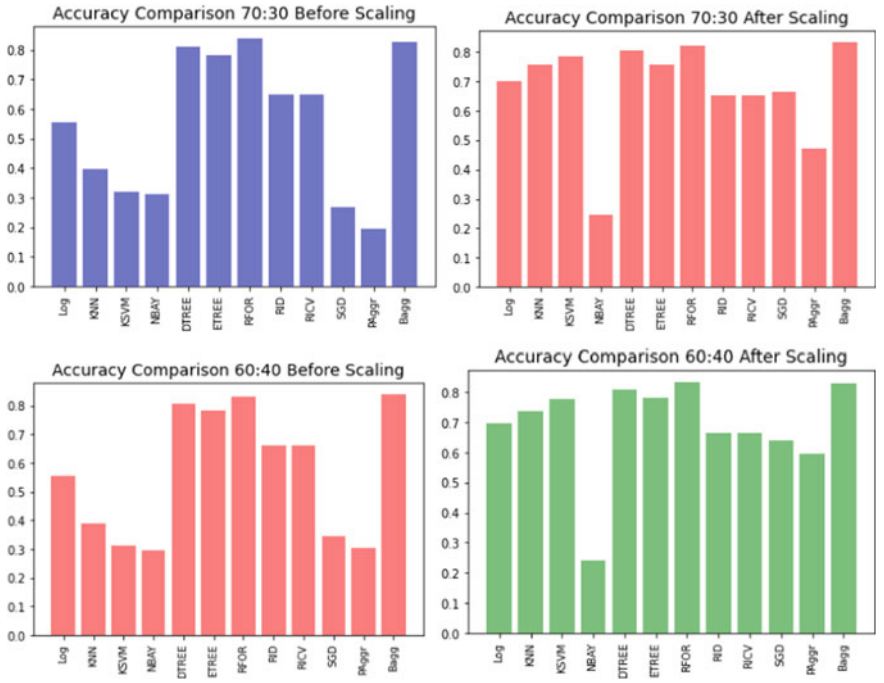


Fig. 4 Accuracy distribution for raw dataset before and after feature scaling

Table 4 LDA classifier performance indices with 80:20 before and after feature scaling

Classifier	Before feature scaling					After feature scaling				
	Precision	Recall	F-Score	Accu	Run Time	Precision	Recall	F-Score	Accu	Run Time
LReg	0.69	0.67	0.66	0.67	0.03	0.69	0.67	0.66	0.67	0.04
KNN	0.80	0.79	0.79	0.79	0.06	0.80	0.79	0.79	0.79	0.05
KSVM	0.82	0.79	0.79	0.79	0.21	0.82	0.79	0.79	0.79	0.23
GNB	0.72	0.71	0.70	0.71	0.00	0.72	0.71	0.70	0.71	0.00
Dtree	0.74	0.74	0.74	0.74	0.02	0.74	0.74	0.74	0.74	0.03
Etree	0.75	0.74	0.74	0.74	0.00	0.75	0.74	0.74	0.74	0.01
RFor	0.95	0.96	0.96	0.96	0.16	0.95	0.96	0.96	0.96	0.14
Ridge	0.68	0.65	0.63	0.65	0.00	0.68	0.65	0.63	0.65	0.00
RCV	0.68	0.65	0.63	0.65	0.00	0.68	0.65	0.63	0.65	0.00
SGD	0.68	0.68	0.67	0.68	0.06	0.69	0.68	0.67	0.68	0.05
PAg	0.35	0.40	0.36	0.40	0.02	0.56	0.55	0.51	0.55	0.01
Bagg	0.79	0.78	0.78	0.78	0.14	0.78	0.78	0.78	0.78	0.14

Table 5 LDA classifier performance indices with 70:30 before and after feature scaling

Classifier	Before feature scaling					After feature scaling				
	Precision	Recall	F-Score	Accu	Run Time	Precision	Recall	F-Score	Accu	Run Time
LReg	0.69	0.67	0.66	0.67	0.02	0.69	0.67	0.66	0.67	0.04
KNN	0.80	0.79	0.79	0.79	0.06	0.80	0.79	0.79	0.79	0.09
KSVM	0.81	0.79	0.79	0.79	0.19	0.81	0.79	0.79	0.79	0.17
GNB	0.71	0.70	0.69	0.70	0.00	0.71	0.70	0.69	0.70	0.00
Dtree	0.74	0.74	0.74	0.74	0.03	0.74	0.74	0.74	0.74	0.02
Etree	0.75	0.74	0.74	0.74	0.00	0.75	0.74	0.74	0.74	0.01
RFor	0.95	0.96	0.96	0.96	0.13	0.95	0.96	0.96	0.96	0.13
Ridge	0.68	0.65	0.63	0.65	0.00	0.68	0.65	0.63	0.65	0.00
RCV	0.68	0.65	0.63	0.65	0.00	0.68	0.65	0.63	0.65	0.01
SGD	0.66	0.64	0.61	0.64	0.05	0.66	0.67	0.66	0.67	0.03
PAg	0.65	0.63	0.62	0.63	0.01	0.53	0.55	0.49	0.55	0.01
Bagg	0.77	0.77	0.77	0.77	0.14	0.78	0.78	0.78	0.78	0.13

$$x_i = U \rightarrow \{0, 1, 2, \ldots n\} \tag{2}$$

After encoding the categorical features in the dataset, the missing values estimation is done by finding the mean of all the values for that particular feature as shown in Eq. (3).

Table 6 LDA classifier performance indices with 60:40 before and after feature scaling

Classifier	Before feature scaling					After feature scaling				
	Precision	Recall	F-Score	Accu	Run Time	Precision	Recall	F-Score	Accu	Run Time
LReg	0.69	0.68	0.68	0.68	0.02	0.69	0.68	0.68	0.68	0.04
KNN	0.79	0.78	0.78	0.78	0.09	0.79	0.78	0.78	0.78	0.11
KSVM	0.80	0.78	0.78	0.78	0.17	0.80	0.78	0.78	0.78	0.16
GNB	0.71	0.70	0.70	0.70	0.00	0.71	0.70	0.70	0.70	0.00
Dtree	0.73	0.73	0.73	0.73	0.02	0.73	0.73	0.73	0.73	0.02
Etree	0.68	0.68	0.68	0.68	0.00	0.68	0.68	0.68	0.68	0.00
RFor	0.95	0.96	0.96	0.96	0.11	0.95	0.96	0.96	0.96	0.10
Ridge	0.68	0.66	0.64	0.66	0.01	0.68	0.66	0.64	0.66	0.00
RCV	0.68	0.66	0.65	0.66	0.00	0.68	0.66	0.65	0.66	0.00
SGD	0.72	0.67	0.64	0.67	0.05	0.65	0.64	0.64	0.64	0.02
PAg	0.53	0.60	0.55	0.60	0.01	0.58	0.52	0.50	0.52	0.01
Bagg	0.77	0.77	0.77	0.77	0.12	0.77	0.77	0.77	0.77	0.12

$$xR_{ij} = \frac{1}{n}\sum_{k=1}^{m}\left(xR_{ij}\right)^{k} \tag{3}$$

The variable "R" denotes the number of rows in the dataset. The variables $\{i, j\}$ represents the instances of the rows and columns, respectively. The pre-processed dataset is now done with feature scaling with min–max normalization with the range values between 0 and 1 and is in Eq. (4).

$$x_i' = \frac{0 + (x_i - \min(x_i))(1 - 0)}{\max(x_i) - \min(x_i)} \tag{4}$$

6 Conclusion

This paper explores the performance of classifying the damage grade for the various distribution of training and testing dataset with 80:20, 70:30 and 60:40 data cross validation. This cross sectioning of the training and testing dataset is also subjected to linear discriminant analysis with 10 components and the results are examined with high accuracy. Implementation results portray shows that the random forest classifier is exhibiting 92% of accuracy after feature scaling and giving 96% of accuracy with 10 component LDA reduced dataset before and after feature scaling. The future enhancement of this paper is to perform principal component analysis for the feature reduction towards the damage grade prediction.

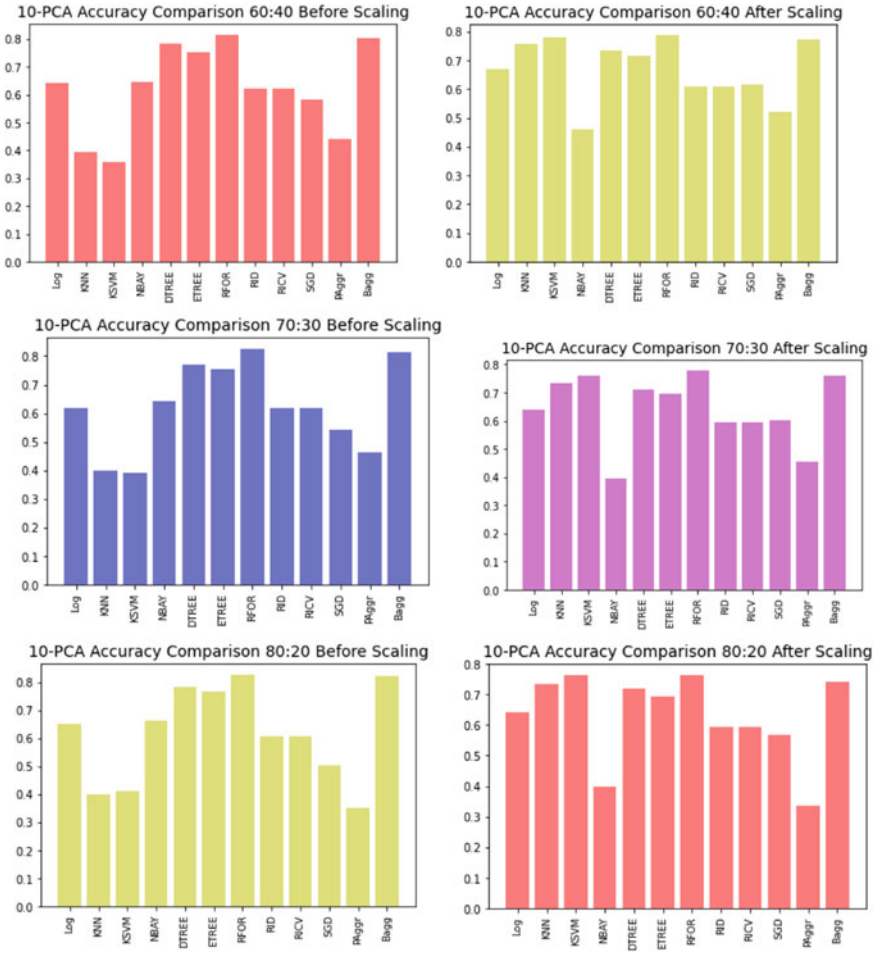


Fig. 5 Accuracy comparison for 10 component LDA reduced dataset

References

1. Shibata A, Motosaka M, Hagiwara Y (2000) Prediction of damage rate of building groups in urban areas by strong earthquakes. In: Conference ICWEE 2000, vol 1, no 1, pp 1369
2. Nakano Y, Maeda M, Kuramoto H, Murakami M (2004) Guideline for post-earthquake damage evaluation and rehabilitation of RC buildings in Japan. In: 13th world conference on earthquake engineering, vol 1, no 1, pp 124
3. Menderes A, Erener A, Sarp G (2015) Automatic detection of damaged buildings after earthquake hazard by using remote sensing and information technologies. In: World multidisciplinary earth sciences symposium, WMESS 2015, vol 15, pp 257–262
4. Korkmaz KA, Abualkibash M (2018) Earthquake damage detection using before and after earthquake satellite images. In: 2018 IEEE international conference on electro/information technology (2018)

5. Rao MVV, Chaparala A (2021) A machine learning-based damage prediction techniques for structural health monitoring. Turk J Comput Math Edu 12(2):3392–3405 (2021)
6. Campos A, Cardona OD, Campos ML (2004) New postearthquake building damage evaluation procedures after recent earthquakes in Colombia. In: 13th world conference on earthquake engineering, pp 2849
7. Yucemen MS, Askan A (2003) Estimation of earthquake damage probabilities for reinforced concrete buildings. In: Seismic assessment and rehabilitation of existing buildings, NATO science series, vol 29. Springer, pp 149–164
8. Ci T, Liu Z, Wang Y (2019) Assessment of the degree of building damage caused by disaster using convolutional neural networks in combination with ordinal regression. Remote Sens 11(23):1234–1248
9. Roeslin S, Ma Q, García HJ (2020) Damage assessment on buildings following the 19th September 2017 Puebla, Mexico earthquake. Earthquake Reconnaissance Build Risk Resilience Evid Base 22(12):1489–99
10. Zheng XW, Li HN, Yang YB, Li G, Huo LS, Liu Y (2019) Damage risk assessment of a high-rise building against multihazard of earthquake and strong wind with recorded data. Eng Struct 200:1096971

Do Age and Education Affect Self-schema? An Exploratory Study Among Individuals

Arti Chandani, N. Srividya, B. Neeraja, and Aditya Srinivas

Abstract Self-schema is known as the dispositions or opinions about him or herself that are usually extracted by self-talk. The schemas are generally made in the childhood. When the child starts looking at the world, at first his parents influence his opinions. The talks and the ideas are imbibed into the person as a child form like schema. For example, the characteristics of good child and a bad child are learnt from parents and near family only. While analyzing the labels of self-aspect, it was found that the individuals who had compartmentalized organization tend to define the self-aspects in narrow and negative terms (Showers in J Pers Soc Psychol 62:1036, 1992, [1]). This paper is based on research done taking few parameters about various self-schemas taking a random sample of 528 using an instrument questionnaire for data collection. As self-scheme is highly subjective, a vast subject few parameters are considered under self-schema. These parameters are tested against the demographical features such as age and education of the respondents. Multivariate analysis is done to analyze if there is any interaction between the demographical features of a person with his self-schema.

Keywords Self-schema · Role performance · Demographical features · Self-image and self-esteem · Body image

1 Introduction

As a baby grows, he keeps observing his immediate family members, neighborhood and a later stage his class, school mates, teachers and others and develops self-talk. These at a later stage form like self-schema. These self-schemas act as predictive agents to one's behavior. Pamela et al. studied a stress model for vulnerability toward stress leading depression along with the self-structured schemas' interactions with negative events in life. Though these are very early contributions in1990s, much research is done in later years.

A. Chandani (✉) · N. Srividya · B. Neeraja · A. Srinivas
SIMS, Symbiosis International University, Pune, MH 411028, India
e-mail: arti.chandani@sims.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_59

2 Review of Literature

The origins of the self-esteem concepts go into as early as 1938. Bierstedt [2] in his article, The Means-End Schema in Sociological Theory, discusses various theories about human being. In the theories, he mentions about positivistic tradition and sociological theory. He also talks about voluntarist theories. He examines means–end schema. Mathew et al. [3] and Bhutani et al. [4] have mentioned in their papers "Perfectionism, Interpersonal Sensitivity, Dysfunctional Beliefs, and Automatic Thoughts" and "Teasing experiences, interpersonal sensitivity, self-schema and psychological distress in youth: An exploratory study" that self-schemas have been discussed by Kaur [5] in her paper "Self-schema in Indian context–Development of a tool".

Later, Rosenberg [6] discussed the self-esteem in four chapters in the book. They had conducted a study on five thousand students of high school age with various national, religious and social backgrounds. The main aim was to study the family effects, neighborhood, groups of minorities, etc., on the self-image of the students and their response to the society.

Systematic review is answering specific questions basing on search strategy with inclusions and exclusions [7]. As mentioned above, thousands of papers are identified in several fields including medical, psychology, etc. But, the papers from 2010s are considered because of the rapid increase in the research in the field. Though several papers are available, papers related to management are quite a few. Though they are very old, they are considered as they belong to the topic. Few latest papers are also listed though they are in medical terms.

Huberman and Chivers [8] study about the sexual self-schemas. A pilot study is conducted to develop a novel measure womens' sexual schema. Chia-Kueil et al. conducted a study on alcohol use and its problems in adolescents. A test was conducted to check cognitive mechanisms in relation to self-schema and self-efficacy of the respondents. Horowitz [9] studies two new concepts called "self" and "identity". This research assesses the organizing self in various levels in a person and helps him to go for higher levels.

Ng [10] conducted a research conducted by two studies. First, a study is conducted on Chinese high school students. Their self-schemas in studying mathematics and various methods of learning are analyzed. Second, this research is extended to a culturally different group for Australian students. Basing on these two researches, a theoretical formulation is developed for both positive and negative schematic students.

Though there is sufficient literature about the self-schema, those are majorly oriented toward medical terms like modification or rectification. Moreover, no paper is devoted toward the effect of demographical features on the self-schema. This paper tries to check the relevance of self-schema in work culture especially in role performance.

The following hypotheses are framed for the study.

H01: There is no association between age and qualification on the self-image variables.
H02: There is no association between age and qualification on self-esteem variables.
H03: There is no association between age and qualification on ideal self-variables.
H04: There is no association between age and qualification on body image variables.
H05: There is no association between age and qualification on role performance.

3 Methods

The scope of the research may be extended to all the employees in almost all the sectors of corporate world and to academic fraternity in understanding the self-schema, thereby the personality of a person. Besides, the corporate world also can gain some insights from their personality while planning their career. The objectives of the study are to figure out the major components of self-schema that really matter and to find out the relation between the self-schema and the age and education of the person.

Random sampling with 528 sample size is adopted for the study, and descriptive analysis, correlations and multivariate analysis are done. SPSS is used for the analysis purpose. The limitations are that the result is limited only to this data set only, the data collected from the questionnaire only is analyzed for results, and finally the number of factors is considered for self-schema. This can be further scaled up to hack or validate the value or belief system of a person. There is another scope for the study where the larger sample can be covered. This research has a very high value in medical field for correction of the beliefs about oneself. There is yet another scope for future study where the sector-specific studies can be done taking the respondents from a particular sample (Fig. 1).

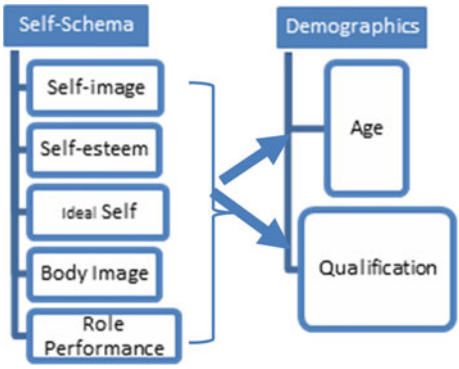


Fig. 1 Proposed model

4 Results

The Cronbach's alpha is 0.909 which indicates that there is high internal validity and reliability in the data (Table 1).

Table 1 Descriptive statistics

	Mean	Std. deviation	*N*
Age	2.21	0.499	528
Qualification	2.10	0.421	528

Multivariate analysis checks the difference in two or more vectors of means. The assumptions for multivariate analysis are all met for both independent and dependent variables. Levene's test for equality of error variances for all the dependent variables is done, and the Sig. *p*-value 0.000 is less than 0.05. Hence, there is homogeneity of variance–covariance matrices for all the variables (Table 2).

There is a statistically significant interaction effect seen between age and educational qualification on the combined self-image variables, $F = 5.370$, $p = 0.000$; Wilks' Lambda = 0.960 (Table 2).

There is a statistically significant interaction effect seen between age and educational qualification on the combined self-esteem variables, $F = 0.288$, $p = 0.000$; Wilks' Lambda = 0.997 (Table 3).

There is a statistically significant interaction effect between age and educational qualification on the combined ideal self-variables, $F = 8.390$, $p = 0.000$; Wilks' Lambda = 0.885 (Table 4).

There is a statistically significant interaction effect observed between age and educational qualification on the combined body image variables, $F = 4.969$, $p = 0.000$; Wilks' Lambda = 0.981 (Table 5).

There is a statistically significant interaction effect observed between age and educational qualification on the combined role performance variables, $F = 8.538$, $p = 0.000$; Wilks' Lambda = 0.883 (Table 6).

Table 2 Multivariate tests for the dependent variable self-image of the respondents

Multivariate tests									
Effect type		Value	*F* value	d*f* for hypothesis	d*f* for error	Sig	Effect type	Value	*F* value
Intercept	Value of Pillai's Trace	0.815	572.455b	4	519	0.000	0.815	2289.820	1
	Value of Wilks' Lambda	0.185	572.455b	4	519	0.000	0.815	2289.820	1
	Value of Hotelling's Trace	4.412	572.455b	4	519	0.000	0.815	2289.820	1
	Value of Roy's Largest Root	4.412	572.455b	4	519	0.000	0.815	2289.820	1
Age	Value of Pillai's Trace	0.033	4.403b	4	519	0.002	0.033	17.611	0.936
	Value of Wilks' Lambda	0.967	4.403b	4	519	0.002	0.033	17.611	0.936
	Value of Hotelling's Trace	0.034	4.403b	4	519	0.002	0.033	17.611	0.936
	Value of Roy's Largest Root	0.034	4.403b	4	519	0.002	0.033	17.611	0.936
Qualification	Value of Pillai's Trace	0.101	14.646b	4	519	0.000	0.101	58.582	1
	Value of Wilks' Lambda	0.899	14.646b	4	519	0.000	0.101	58.582	1
	Value of Hotelling's Trace	0.113	14.646b	4	519	0.000	0.101	58.582	1
	Value of Roy's Largest Root	0.113	14.646b	4	519	0.000	0.101	58.582	1
Age * Qualification	Value of Pillai's Trace	0.040	5.370b	4	519	0.000	0.040	21.481	0.973
	Value of Wilks' Lambda	**0.960**	**5.370b**	4	519	0.000	0.040	21.481	0.973
	Value of Hotelling's Trace	0.041	5.370b	4	519	0.000	0.040	21.481	0.973
	Value of Roy's Largest Root	0.041	5.370b	4	519	0.000	0.040	21.481	0.973

Table 3 Multivariate tests for the dependent variable self-esteem of the respondents

Effect type		Value	*F* value	d*f* for hypothesis	d*f* for error	Sig
Intercept	Value of Pillai's Trace	0.849	583.105	5	518	0
	Value of Wilks' Lambda	0.151	583.105	5	518	0
	Value of Hotelling's Trace	5.628	583.105	5	518	0
	Value of Roy's Largest Root	5.628	583.105	5	518	0
	Value of Pillai's Trace	0.068	7.561	5	518	0
	Value of Wilks' Lambda	0.932	7.561	5	518	0
	Value of Hotelling's Trace	0.073	7.561	5	518	0
	Value of Roy's Largest Root	0.073	7.561	5	518	0
Qualification	Value of Pillai's Trace	0.009	0.899	5	518	0.482
	Value of Wilks' Lambda	0.991	0.899	5	518	0.482
	Value of Hotelling's Trace	0.009	0.899	5	518	0.482
	Value of Roy's Largest Root	0.009	0.899	5	518	0.482
Age * Qualification	Value of Pillai's Trace	0.003	0.288	5	518	0.920
	Value of Wilks' Lambda	**0.997**	**0.288**	5	518	0.920
	Value of Hotelling's Trace	0.003	0.288	5	518	0.920
	Value of Roy's Largest Root	0.003	0.288	5	518	0.920

Table 4 Multivariate tests for the dependent variable ideal self of the respondents

Effect type		Value	*F* value	d*f* for hypothesis	d*f* for error	Sig
Intercept	Value of Pillai's Trace	0.901	584.347	8	515	0
	Value of Wilks' Lambda	0.099	584.347	8	515	0
	Value of Hotelling's Trace	9.077	584.347	8	515	0
	Value of Roy's Largest Root	9.077	584.347	8	515	0
Age	Value of Pillai's Trace	0.066	4.561	8	515	0
	Value of Wilks' Lambda	0.934	4.561	8	515	0
	Value of Hotelling's Trace	0.071	4.561	8	515	0
	Value of Roy's Largest Root	0.071	4.561	8	515	0
Qualification	Value of Pillai's Trace	0.104	7.449	8	515	0
	Value of Wilks' Lambda	0.896	7.449	8	515	0
	Value of Hotelling's Trace	0.116	7.449	8	515	0
	Value of Roy's Largest Root	0.116	7.449	8	515	0
Age * Qualification	Value of Pillai's Trace	0.115	8.390	8	515	0
	Value of Wilks' Lambda	**0.885**	**8.390**	8	515	0
	Value of Hotelling's Trace	0.130	8.390	8	515	0
	Value of Roy's Largest Root	0.130	8.390	8	515	0

Table 5 Multivariate tests for the dependent variable body image of the respondents

Multivariate tests						
Effect type		Value	*F* value	d*f* for hypothesis	d*f* for error	Sig
Intercept	Value of Pillai's Trace	0.728	697.213b	2	521	0.000
	Value of Wilks' Lambda	0.272	697.213b	2	521	0.000
	Value of Hotelling's Trace	2.676	697.213b	2	521	0.000
	Value of Roy's Largest Root	2.676	697.213b	2	521	0.000
Age	Value of Pillai's Trace	0.022	5.852b	2	521	0.003
	Value of Wilks' Lambda	0.978	5.852b	2	521	0.003
	Value of Hotelling's Trace	0.022	5.852b	2	521	0.003
	Value of Roy's Largest Root	0.022	5.852b	2	521	0.003
Qualification	Value of Pillai's Trace	0.006	1.596b	2	521	0.204
	Value of Wilks' Lambda	0.994	1.596b	2	521	0.204
	Value of Hotelling's Trace	0.006	1.596b	2	521	0.204
	Value of Roy's Largest Root	0.006	1.596b	2	521	0.204
Age * Qualification	Value of Pillai's Trace	0.019	4.969b	2	521	0.007
	Value of Wilks' Lambda	**0.981**	**4.969b**	2	521	0.007
	Value of Hotelling's Trace	0.019	4.969b	2	521	0.007
	Value of Roy's Largest Root	0.019	4.969b	2	521	0.007

Table 6 Multivariate tests for the dependent variable role performance of the respondents

Multivariate tests						
Effect type		Value	*F* value	d*f* for Hypothesis	d*f* for Error	Sig
Intercept	Value of Pillai's Trace	0.910	653.638b	8	515	0.000
	Value of Wilks' Lambda	0.090	653.638b	8	515	0.000
	Value of Hotelling's Trace	10.154	653.638b	8	515	0.000
	Value of Roy's Largest Root	10.154	653.638b	8	515	0.000
Age	Value of Pillai's Trace	0.149	11.251b	8	515	0.000
	Value of Wilks' Lambda	0.851	11.251b	8	515	0.000
	Value of Hotelling's Trace	0.175	11.251b	8	515	0.000
	Value of Roy's Largest Root	0.175	11.251b	8	515	0.000
Qualification	Value of Pillai's Trace	0.146	10.979b	8	515	0.000
	Value of Wilks' Lambda	0.854	10.979b	8	515	0.000
	Value of Hotelling's Trace	0.171	10.979b	8	515	0.000
	Value of Roy's Largest Root	0.171	10.979b	8	515	0.000
Age * Qualification	Value of Pillai's Trace	0.117	8.538b	8	515	0.000
	Value of Wilks' Lambda	**0.883**	**8.538b**	8	515	0.000
	Value of Hotelling's Trace	0.133	8.538b	8	515	0.000
	Value of Roy's Largest Root	0.133	8.538b	8	515	0.000

5 Discussions

Body image can be considered as on internalized appearance which leads to the behavior of individuals. Appearance of self-schema and discrepancy self-mediated the effect of use of Instagram on satisfaction of body among students of university who were using Instagram [11]. The response of women to media images varied due to differences in individual variables, appearance self-scheme and cognitive representation of organized information about the self vis-a-vis to appearance (Jung and Lennon [12]. It was reported by [13] that the self-schema appearance is able to explain the image of body better than the appearance self-discrepancy. Lambert and Wedell [14] presented a theoretical framework for predicting effects of self-knowledge on judgment of others. The respondents having high complexity could attend and encode the feedback which was disconfirming, while the respondents, who had low complexity, rejected the feedback and reasserted the positives aspect of the self.

References

1. Showers C (1992) Compartmentalization of positive and negative self-knowledge: keeping bad apples out of the bunch. J Pers Soc Psychol 62(6):1036
2. Bierstedt R (1938) The meanings of culture. Philos Sci 5(2):204–216
3. Mathew M, Sudhir PM, Mariamma P (2014) Perfectionism, interpersonal sensitivity, dysfunctional beliefs, and automatic thoughts: the temporal stability of personality and cognitive factors in depression. Int J Ment Health 43(1):50–72
4. Bhutani R, Sudhir PM, Philip M (2014) Teasing experiences, interpersonal sensitivity, self-schema and psychological distress in youth: an exploratory study. Psychol Stud 59(3):241–251
5. Kaur L (1957) Self-schema in Indian context–development of a tool. Unpublished M. Phil dissertation Bangalore. In: Shah A, Rao K (eds) Psychological research in mental health and neurosciences, vol 2007, p 372–373
6. Rosenberg M (2015) Society and the adolescent self-image. Princeton University Press
7. Gough D, Oliver S, Thomas J (eds) (2017) An introduction to systematic reviews. Sage
8. Huberman JS, Chivers ML (2020) Development of a novel measure of women's sexual self-schema organization☆. Person Indiv Differ 155:109665
9. Horowitz MJ (2012) Self-identity theory and research methods. J Res Pract 8(2):M14–M14
10. Ng CH (2005) Academic self-schemas and their self-congruent learning patterns: findings verified with culturally different samples. Soc Psychol Educ 8(3):303–328
11. Ahadzadeh AS, Sharif SP, Ong FS (2017) Self-schema and self-discrepancy mediate the influence of Instagram usage on body image satisfaction among youth. Comput Hum Behav 68:8–16
12. Jung J, Lennon SJ (2003) Body image, appearance self-schema, and media images. Fam Consum Sci Res J 32(1):27–51
13. Jung J, Lennon SJ, Rudd NA (2001) Self-schema or self-discrepancy? Which best explains body image? Cloth Text Res J 19(4):171–184
14. Lambert AJ, Wedell DH (1991) The self and social judgment: effects of affective reaction and own position on judgments of unambiguous and ambiguous information about others. J Pers Soc Psychol 61(6):884

Automatic Waste Segregation System: An IOT Approach

Siddharth Nair, C. K. M. Ganesan, A. Ram Gunasekaran, S. Sabarish, and S. Asha

Abstract Effective waste management and disposal are major issues in today's world, especially in countries like India where a majority of individuals do not use separate trash bins for recyclable and non-recyclable waste. This is a major cause of pollution and diseases in the country. However, separating them manually is a tedious process and also not efficient. Improper segregation results in pollution and affects the environment. We have developed an automated trash bin which autonomously classifies and segregates wastes as biodegradable and non-biodegradable and places them in the respective bin. We have achieved the same using IOT by passing the image data to the cloud to perform analytics and return the classified output to the controller of the bin and trigger it to place the waste in the correct container. We also send the usage data to the cloud to perform analytics about the bin usage and deploy it in a Web app to report the live status.

Keywords Machine Learning · Waste segregation · Raspberry Pi · IOT · Convolutional Neural Network · Cloud · Analytics

1 Introduction

Waste segregation is included in the law because it is very easy to recycle. Efficient classification of waste means that less waste goes to landfills which makes it cheaper and better for people and the environment. It is also important to segregate waste for public health. In particular, hazardous waste can cause long-term health problems, so it is very important that it is disposed off properly and safely and that it is not mixed with ordinary waste from your home or office. A large amount of people suffer from

S. Nair (✉) · C. K. M. Ganesan · A. Ram Gunasekaran · S. Sabarish · S. Asha
Vellore Institute of Technology Chennai Campus, Chennai, Tamil Nadu 600127, India
e-mail: siddharthmani2000@gmail.com

S. Asha
e-mail: asha.s@vit.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_60

illnesses every day because of the effects of unhealthy waste disposal system in India. It is also a major cause of global warming which further degrades the environment.

For effective waste management, segregation has to be encouraged in urban areas to reduce landfills gradually so that eventually, there are no more landfills in the future. Even in Western Europe where they have been working on building scientific waste management systems for the past 30 years, almost one-third of the waste still continues to be dumped in landfills.

Proper waste segregation not only protects the environment but also provides the earth with resources; for example, biodegradable wastes can be used as a source of fuel for biogas, which is not utilized properly when waste segregation is not done. The creation of these renewable "green" energy resources can substantially reduce the toll being taken on fossil fuels and other conventional non-renewable resources. It can be used to generate electricity in various rural areas as well instead of deploying dangerous and health degrading thermal and nuclear power plants.

It is inhumane to expect the government and corporation officials to segregate mixed recyclable and non-recyclable waste by hand. In order to optimize the separation process, the ideal way is to segregate them at source so that the different kinds of waste can be collected and handled separately without any extra hassle. Therefore, there is a need for a system which automatically segregates wastes for the people at the home so that they need not perform the task manually.

According to Wei et al. [1], convolutional neural network (CNN) has provided high accuracy in single-label image classification-related tasks. However, in the case of multi-label image classification, it does not perform as good because of complex object layouts and lack of adequate training images. In this work, the authors have proposed a CNN architecture that is flexible called as HCP (Hypotheses-CNN-Pooling), where a number of object segments are provided as inputs; then, this data is passed to a shared CNN network, and results from the different modules of the shared CNN are aggregated to classify the image to its appropriate label. The characteristics of this CNN architecture is: (1) easier to train as no information is required; (2) robust to noisy and/or redundant hypotheses; (3) model can be pre-trained with a custom dataset; and (4) provides better results on multi-label classification.

In another paper, Qin et al. [2], the authors address that with the increase in biological images, classifying them becomes challenging, and they suggest that CNN provides a solution for this problem. The paper suggests a biological image classification algorithm using an adjusted CNN which improves the accuracy of classification and also makes the model more lightweight. In this paper, the input images to the CNN model are resized to a large size, and the modules are replaced with an inverted residual block module as it has fewer computational cost and parameters. Upon thoroughly analyzing the computational costs and classification accuracy of the proposed method using five well-known benchmark datasets, the results show that when compared with existing image classification methods, the proposed method shows better accuracy and reduces network parameters and computational costs.

Paper Pamintuan et al. [3] state that one of the problems of increasing world population is environmental hazards. This can be attributed to increased waste generation in every household. In Philippines, it has reported that around 40,087 tons of waste

are being generated per day as of 2016. If not disposed properly, these wastes can pollute the water bodies around the area and cause major environmental problems. To mitigate these problems, it calls for a better waste disposal system. Nowadays, where everything is digitalized and smart city being a vision of many nations, it would be effective to implement Internet of Things (IOT) in the waste disposal system as a first step to realize it. Data can be collected by implementing IOT technology, which can be analyzed further to improve the process. Thus, a trash bin with sensors in it comes as an effective approach. It can classify a given waste as either biodegradable or non-biodegradable, thus minimizing human intervention. It has a considerable accuracy of 99% on trained datasets and 75% for unknown images. This can be improved further using the data collected and training it with a more images. The background of the image has less interference in the classification, thus making it practical for real-world application.

According to Azis et al. [4], recycling of waste from households and industries is one of the methods that has been proposed to reduce the ever-increasing pressure on landfills. As different waste types require different techniques to dispose, proper segregation based on types is mandatory for efficient recycling. It is not efficient to separate wastes manually based on the type. The authors have developed a solution based on deep learning and computer vision concepts using which wastes can be classified into six different categories—glass, plastic, metal, cardboard, paper and others. The Inception V3 model (based on convolutional neural networks) is used for the classification with a trained dataset obtained from the Internet. High classification accuracy of 92.5% is achievable using the proposed method. The authors say that this method would help in the automating the process of waste segregation, therefore reducing human involvement.

2 Proposed Methodology

When the user places the waste in the bin, the load balancer detects a voltage fluctuation, and this is amplified by the HX711 module so that the weight of the trash can be detected. On sensing a weight change, the Raspberry Pi triggers the camera module to take a picture of the trash placed on top of the bin. After capturing the image, the Raspberry Pi sends the image to the cloud using HTTP post request for waste classification as biodegradable or non-biodegradable. The Raspberry Pi also sends the timestamp, binary image data, and the weight detected to Azure Cosmos DB for further analytics and visualizations so that the user can view log data of the trash bin. The convolutional neural network-based image classification model is deployed in the cloud. Through the aforementioned HTTP post request, it receives the image and the weight of the trash placed and classifies it as biodegradable or non-biodegradable and sends the response to the Raspberry Pi module. Heroku cloud is also used to deploy the Web portal for smart sort to make it more accessible. This makes sure that data is received from the database in real time so that trash logs along with visualizations are present to understand bin activity. Once the image of the waste is classified

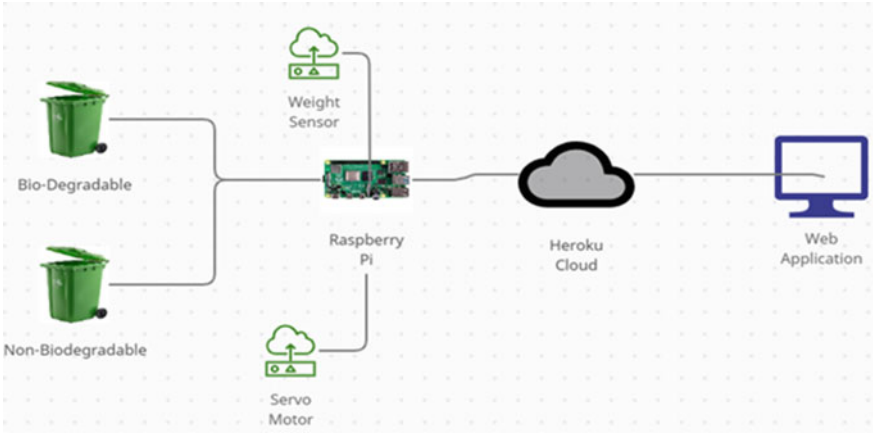


Fig. 1 Overall architecture of the system

and the result is sent back to the Raspberry Pi module, it triggers the servo motors to push the waste into their respective bins. For example, if the waste is biodegradable, the motor will be triggered to push the waste into a green bin. On the other hand, non-biodegradable waste is pushed into a blue bin by the servo motor. The trash log data is displayed in a dashboard using a Web-UI. This website contains details of the live usage of the trash can along with visualizations performed. The data sensed and collected by the Raspberry Pi module is also sent via a serial connection so that the trash data will be live on the dashboard. The methodology depicted is illustrated in Fig. 1.

2.1 Preprocessing

Input images are standardized for uniformity by performing normalization methods. This is performed using OpenCV in Python. The images are processed to suppress unwanted distortions and noise from the image. It is also present to enhance the image features so that optimum results are procured in text recognition and extraction. Initially, the colored input images are resized to 500x400. Noise is removed from the resultant image using the non-local means denoising algorithm, which replaces each pixel value with the average of the selection of other surrounding pixel values, and the average is performed only for regions that have patches nearby to the current spot [5]. As a result, this algorithm can produce fine-textures, which will get blurred by other denoising algorithms. Next, the contrast of the image is normalized for image segmentation. Finally, they are converted into greyscale to improve feature extraction for the subsequent stages.

2.2 Camera

A camera is used to capture an image of the surroundings. In our application, the Raspberry Pi triggers the camera to take a picture of the trash that is placed on top of the waste bin. This image will be stored locally and sent to the cloud for processing

2.3 Servomotor

A servomotor is a rotary or linear actuator that can control angular or linear position, velocity, and acceleration with precision. It is made comprised of an appropriate motor and a position feedback sensor. It also necessitates a complex controller, which is frequently a separate module created exclusively for servomotors. In our system, servomotors are used to push the classified waste into their respective bins. Once the class of the waste is received from the cloud, it acts as an actuator so that the waste can automatically be segregated. Although the term servomotor is frequently used to refer to a motor appropriate for use in a closed-loop control system, it is not a distinct class of motor.

2.4 Load Cell

A load cells are devices that detects the size of a load (either a force or a weight) and transforms the force into a quantifiable electrical output. They serve a variety of functions, including determining the weight of an object, recording weight changes over time, and quantifying strain and tension on a surface [6]. When a force is applied to a load cell, it transforms into a metal core with a set of electrical resistances. However, once the force is released, it reverts to its previous state. The reversibility of this substance determines the load cell's quality and accuracy. With the HX711, it is possible to extract measurable data from a load cell and a strain gauge. This is used by our system to create a voltage change on detecting a weight placed on top of it.

2.5 HX711

When a force is applied to a load cell, it transforms into a metal core with a set of electrical resistances. However, once the force is released, it reverts to its previous state. The reversibility of this substance determines the load cell's quality and accuracy. With the HX711, it is possible to extract measurable data from a load cell and a strain gauge. The HX711 chip combines a regulated output voltage with various

input and output circuits for high integration, quick response, and known for its anti-interference performance. This is interfaced with the load balancing cell so as to amplify the input signal and trigger the Raspberry Pi script.

2.6 Cloud

We have deployed our deep learning model on Heroku cloud. The model receives the image captured by the Raspberry Pi module via http request. The model trained is exported and deployed in the cloud. The model receives the input image from the http request and classifies the waste. After classifying, it sends back the response to the Raspberry Pi module. The Raspberry Pi triggers the servo motor and pushes the waste into its respective bin.

2.7 Azure COSMOS

Cosmos DB is a highly scalable NoSQL database technology that runs on Azure from Microsoft. Cosmos database is a multi-model, multi-query-API database that is horizontally scalable, globally distributed, fully managed, and has minimal latency [7]. We have used this in our system to store the data received from the Raspberry Pi module—the binary data of the image captured, weight of the trash, timestamp in our database which is used for further analytics. The data is extracted from Cosmos DB so that it can be displayed on the Web portal in the form of logs and visualizations.

3 Results

The proposed model worked accurately on the wastes with which we tested. The CNN model classified the wastes with an accuracy of 96.7% on the test dataset. And the hardware part also worked flawlessly into disposing the waste into the respective bin without any error or deterioration of the components. The load balancing cell and HX711 modules provided an accurate calculation of the weights of the waste collected with an accuracy of 98% and an average error rate of +/− 0.5 g. Segregation of biodegradable and non-biodegradable waste is an extremely demeaning and cumbersome task. This creates a need for a system which can automate the entire process so that the waste can be processed more efficiently. This is extremely beneficial with regards to conserving the environment and reducing pollution. It is also essential to effectively converted segregated waste into electrical energy through conversion into biogas.

4 Conclusion

In the future, we intend to develop this prototype into a full-fledged product which would benefit the society and help in waste segregation which is not being done properly in today's world. This product aligns with the Indian Government's Swachh Bharat initiative and can be deployed in major cities so that the corporation authorities will not have to manually segregate waste. The system can be further improved to notify officials when the trash cans are full so that it is not left stagnant and can be collected quickly. Area-wise analytics can be displayed on the dashboard so that the government can identify the locations in need of more trash bins.

References

1. Wei Y et al (2016) HCP: a flexible CNN framework for multi-label image classification. IEEE Trans Pattern Anal Mach Intell 38(9):1901–1907. https://doi.org/10.1109/TPAMI.2015.2491929
2. Qin J et al (2020) A biological image classification method based on improved CNN. Ecol Inform 58:101093
3. Pamintuan M, Mantiquilla SM, Reyes H, Samonte MJ (2019) i-BIN: an intelligent trash bin for automatic waste segregation and monitoring system. In: 2019 IEEE 11th international conference on humanoid, nanotechnology, information technology, communication and control, environment, and management (HNICEM), pp 1–5 https://doi.org/10.1109/HNICEM48295.2019.9072787
4. Azis FA, Suhaimi H, Abas E (2020) Waste classification using convolutional neural network. In: Proceedings of the 2020 2nd international conference on information technology and computer communications, pp 9–13
5. Yang M, Thung G (2016) Classification of trash for recyclability status. CS229 project report 2016
6. Meng X, Tan X, Wang Y, Wen Z, Tao Y, Qian Y (2019) Investigation on decision-making mechanism of residents' household solid waste classification and recycling behaviors. Resour Conserv Recycl 140:224–234
7. Altikat AAAGS, Gulbe A, Altikat S (2021) Intelligent solid waste classification using deep convolutional neural networks. Int J Environ Sci Technol 1–8

Employer Brand for Right Talents in Indian IT Companies

Pravin Kumar Bhoyar and Rajiv Divekar

Abstract Employer branding is a concept which deals with human resource. It has become very important in today's time when the reliance on intellectual capital is more. The concepts started from the ashes of Industrial age in the 1980s and made its presence felt in the 1990s. In post liberalization era, we have seen brand name being established not for the product they deliver, but as a place to work. Employer branding is a process with which the organization places an image of being "**a great place to work**" in the minds of the target future employees. The best place to work is actually feedback given by employees as to what they think of your organization, "a **place to work**." In this research paper, we have highlighted some of the important points which are relevant for the IT companies as far as employer brand is concerned. The research has showed that IT companies have separate budget for employer brand, and they really do not have difficulty in attracting the **right talent** for their organization. These IT companies are well versed with the fact as to where the potential employees hang out and what are the factors which the potential as well as current employees are looking in an organization.

Keywords Employer branding · IT Companies · Indian · Employer

1 Introduction

Employer branding is a concept which deals with human resource. It has become very important in today's time when the reliance on intellectual capital is more [1]. The concepts started from the ashes of Industrial age in the 1980s and made its presence felt in the 1990s. In post liberalization era, we have seen brand names being established not for the product they deliver, but as a place to work.

Employer branding is a process with which the organization places an image of being "**a great place to work**" in the minds of the target future employees. The best

P. K. Bhoyar (✉) · R. Divekar
Symbiosis Institute of Management Studies, Symbiosis International (Deemed University), Pune, MH 411020, India
e-mail: pravink@sims.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_61

place to work is actually a feedback given by employees as to what they think of your organization, "**a place to work**." This information is very useful as organization can bring about the changes and address the employees' issues and concerns and can make their organization as a better place to work. If the organizations are considered to be the better place to work, the employees are more satisfied and motivated [2] which results in having fewer turnovers and less hunt for the organization for talent.

Employer branding uses some of the tools of management like marketing research, public relations, and advertising to create the image the company intends to portray, into the minds of potential employees [3]. Nowadays, "**employer by choice**" has become a generic product. As with many HR practices, employment branding can have a huge impact on the business with a relatively low investment if managed well. However, the reverse is also true. This shows the criticality involved in managing employer branding in an organization.

2 Review of Literature

Simon Barrow (Chairman, People in Business) is considered to be the creator of the term, employer brand. It is a technique which makes the identity of the employer clear, and this result in the companies' image as an employer. A company's employer brand is its "**trust mark,**" i.e., the sum of the company's parts, what makes it unique, and the personality of the company.

Though the concept of employer branding is not very new, but what is happening is that more and more companies are looking to harness the power of their employer brand with a single and sole aim: **Talent.** Employer branding is about the health of organization's talent interaction. According to Hodes global survey [4], understanding the pros and cons of your organization's ability to deliver a great workplace should provide you with all the breadth you need to shout from the rooftops, "**We are a great place to work. Join us**"

The time has changed a lot. There was a time when people used to get hired first and then used to settle in that city, but this trend has revered now. Now, people choose the city first, i.e., where they want to live, and then, they try to get hired in that city or place. Names such as IBM, Sundram, Infosys, and TCS are not known only for the product and services they deliver, but people respect them and take their name with pride. People want to get associated with these companies and want to work with them.

3 The Components of Employer Branding

Employer branding generates a sense of being more caring for your employees which inculcates the feeling of belongingness in the organization. It brings the spirit of cooperation between HR department and marketing department, whether the ultimate

goal is not to sell the product and get profits, but also to sell the employer brand and get a workforce which is loyal, dedicated, and self-motivated.

Employer branding is as good as selling a product. When someone is approaching new clients and consumer, one has to try hard to improve one's own image. Similarly, when trying to attract the "**right talent**" or so-called applicant, one has to see one's image. Here, the market is of labor or workforce, in contrast to consumers or clients [5]. But the importance is same for the company. The only thing that may differ is the timing of the benefit what a product brand may give and what employer branding may give. The benefits of product brand are apparent in a very short time period as the sales figure start increasing, while in the case of benefit of employer, brand may not be visible immediately, but in the long run, it is definitely visible. There are **five components** of employment branding (Fig. 1):

1. **Work life balance**: This component is concerned with how much **travel** a job will require, whether there will be any **flextime** for the workers to choose. It is also concerned with the location of the job and vacations which an employee can avail. Each employee would like to strike a proper balance between work and personal life. So all the employees evaluate the company on this parameter as to how the balance will be there between work and personal life. If the employees feel that by joining in a particular company they can manage both, work and personal life, very well, they will be interested to join in the company.
2. **Company culture**: It is basically concerned with the quality of the senior team, the culture toward problem-solving. It also focuses on the risk-taking environment and the development reputation. The culture of the company plays an important role in retaining as well as attracting the work force. If the culture is where the employees are given responsibility clubbed with authority, they will be motivated and interested. Similarly, if they are allowed to take some bearable amount of risk, it generates a sense of being responsible.
3. **Company Brand Strength**: Reputation. What is the appeal of the company's brand name to its target or potential employees? This is related to as to what image company has in the market place. Can employees take the name of the

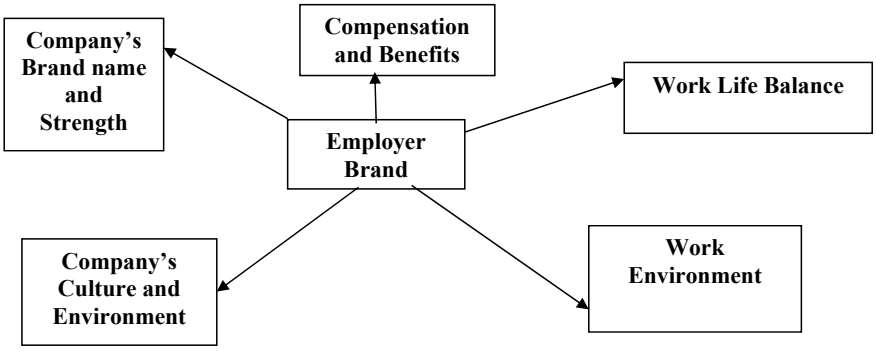


Fig. 1 Five components of employer branding (*Source*: Author's representation)

company with pride? This is also important as employees see the long-term value addition in their own career. If they feel that by joining in a particular company will give them edge or some apparent benefits, they will be more interested to be the part of the company.

4. **Compensation and benefits**: What will be the basic salary, will there be any bonus, will there be any stock option and retirement contribution and health benefits. This is the monetary benefit given to the employees. Actually, this is the first thing, where at times, employees may decide join or not to join the companies. So, the compensation and benefits should be designed keeping this thing in mind. The company should also think as to what others (competitors) are offering for the same type of job. The whole package should be designed keeping the market situation in the mind. Most of the employees who leave organization are attributed to the monetary reasons only.
5. **Work environment**: Manager quality, co-worker quality, recognition, empowerment, work challenge, cutting-edge work, international mobility, role clarity. The work environment plays an important role in retaining and attracting the employees. The company should have an environment which promotes team work, recognition, and prestige for the workers. Company should try to build a good and conducive environment which is beneficial for the employees, and it yields productivity and fewer turnovers for the organization.

4 Employer Branding Process

In order to have a strong employer brand, the company has to do introspection of itself so that it can come out with a strong, differentiated, and sustainable employer brand in the labor market. The company should project its image, what it wants to propagate, in the minds of current and potential employees. The company should design the objective of employer brand as to what the company wants to achieve by resorting to employer brand. There are certain steps which are to be followed before the company can come out with a final product known as employer brand. The steps are as follows: -

1. **Forming the Core Team**: First, the company has to form a core team which consists of individual from several departments so that it is a cohesive group. The individual selected should have willingness to be the part of employment branding process and should be willing to work together. This team will be the leading team for implementing the employment branding process in an organization for which they should be provided with all the necessary and required information. The team members should be expert from the area viz. advertising, marketing, marketing research, public relation, and human resource development.
2. **Building the Good Employer Proposition**: In order to build the strong employer proposition, the company has to see what the strengths of the company are and how to make these strengths stronger. The company should also look

at competitors as to what they are offering in terms of employer branding, and company should try to differentiate its employer brand and try to make somewhat unique in the labor market. An appropriate consideration should be given for the shareholder's expectation from the employer brand image [6].

3. **Benchmarking competition**: With the help of core team, the company should try to find the extent of competition available for recruiting skill set. Here the company should not only look into the extent but also the degree to which company is facing competition to have better idea of its position vis to competitors.

5 Successful Employer Brand

Employer branding is considered to be the DNA of the organization. Employer branding tries to build the good image of the organization in the minds of current and potential employees to make them feel that, "**Yes, this is the best place for me to work**." It binds the employees with the organization and helps in achieving loyalty as well as fewer turnovers. Employer branding is also about the relationship between an employer, customer, and stakeholders. The employer brand has become very important in today's time where the mobility of the labor is very high, and it is of paramount importance to retain the right, "talent culture fit," which has become very difficult in today's time. Employer branding ensures the loyalty, commitment, and the satisfaction on the employees' part which ultimately leads to the growth, profitability of the company [3] **How the employment offers, if appropriately constructed, both establishes and perpetuates the employment brand** (Fig. 2) [5].

6 Benefits of Successful Employer Brand

The company is trying hard to build a strong employer brand with a definite objective of being viewed as, "Preferred **employer**." The employer brand brings about the desired changes in the outlook of the company in the labor market and the long term. Few of the benefits of the successful employer brand are:

1. **Less effort required during recruitment process**: Milward-Brown 2001 Brandz survey found that the organization which had strong employer brand had to work less during recruitment process as compared to those organizations which has week and less strong employer brand. Weak and unfamiliar brands had to work much harder in their recruitment message to build their brand, which is again very difficult process.
2. **Reduced recruitment cost**: If the company is well knowing or known in the labor market, the effort required in recruitment are quite less as the candidates or potential employees will come forward themselves. It reduces the cost of

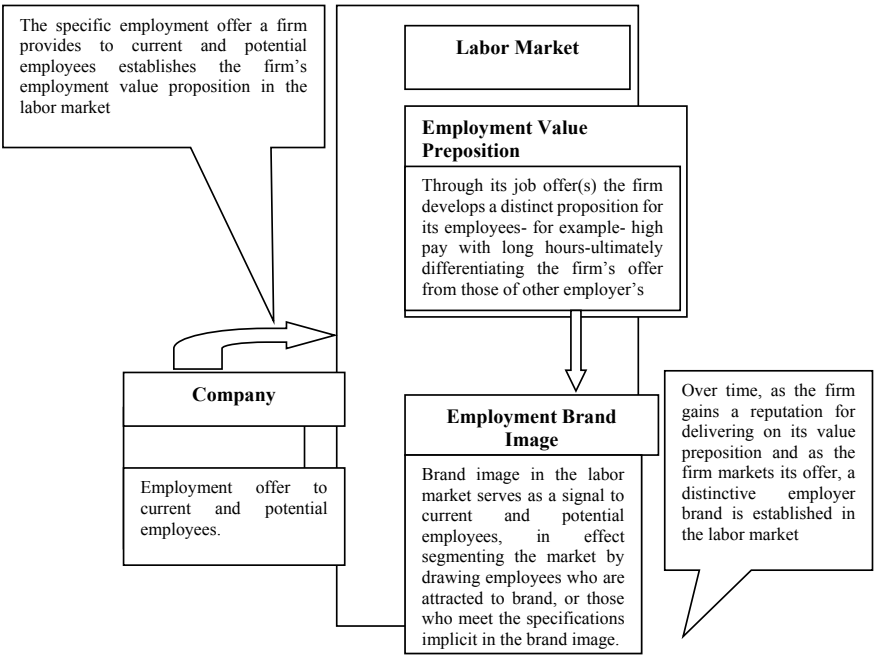


Fig. 2 Importance of successful employer branding for organization (*Source*: Author's representation)

recruitment where the company has to search for the right talent and spend a lot on finding the whereabouts of the talent and then making them to come forward for the organization.

3. **Higher Staff retention:** When company is having well-known employer brand, the staff turnover is quite less which is as good as staff retention. The employer brand brings an array of benefits to the existing employees. Employer brand adds value to the employee's future career, and when it is perceived, employees would like to stay with the company.
4. **Increased profitability:** With the reduction in recruitment cost, which is one of the huge costs, the profitability is bound to go up. Though the benefits of the employer brand may not be perceived and felt immediately, but within a year or two, employer brand will start showing its result. One of the major contributions is the increase in profit.

7 Analysis of the Research

This paper is a research paper, where the objective was to find out the employer branding in Indian IT companies. In order to find out the level of awareness and extent of compliance to employer brand concept, we designed a questionnaire consisting

of 16 questions. Questionnaire was the only method which was used here to find the data about employer brand. All the questions and their respective responses are not analyzed here. Few important questions are analyzed here as only these are more relevant for the topic mentioned here. The respondents were from Scope international, ITC Infotech, HCL, Infosys, Cybernet/support, Keane etc. We gave the questionnaire to 25 respondents. We have analyzed the data and analysis is presented in the following sections (Fig. 3).

This graph is representation of the question number 7, which pertains to whether is there any difficulty in attracting quality staff in their organization. 40% of the respondents were of the view of that they are having difficulty in attracting quality staff in their organization while 60% of the view that there is no difficulty. The 40% responses showed that they had **difficulty in attracting in quality staff** in their organization. This difficulty was attributed to the fact that out of that, 50% of the companies did not follow employer brand to their fullest extent. These 50% companies had employer brand awareness only to the moderate level.

This shows the importance of employer brand in attracting the quality staff. The 60% responses showed that **they had no difficulty in attracting in quality staff** as they had full awareness of employer brand, and they were following employer brand practice in their respective companies. This clearly shows how and what role the employer brand can play in attracting the quality staff in their organizations. There is an increasing trend in the labor market to have a strong employer brand and organizations have started reacting to this demand and ultimately getting benefit out of it. These organizations are increasingly becoming the, "**Best places to work or preferred place for work**." Once the organization has established its employer brand, then it will not have any difficulty in attracting the right talent [2]. Furthermore, employer brand is also beneficial for the organizations as satisfied employees are the assets for the organizations, and these satisfied employees are self-motivated and results in fewer turnovers (Fig. 4).

The above chart is the graphical representation of the question number 5 which relates to the amount of budget for employer brand, if any. All the respondents had a separate budget for employer brand, but the quantum of the budget differed among respondents (companies). 5 respondents (20%) had an annual budget of less than Rs. 1, 00,000, while 20 respondents had (80%) more than Rs. 5, 00,000 as annual

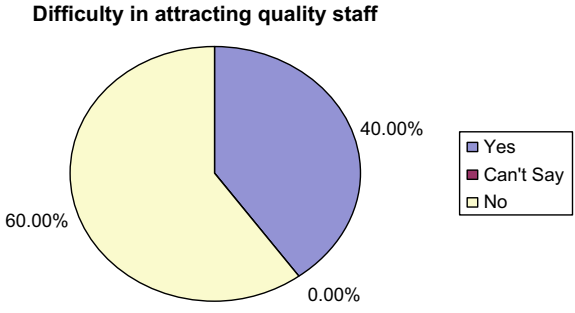


Fig. 3 Difficulty in attracting quality staff (*Source*: Author's representation)

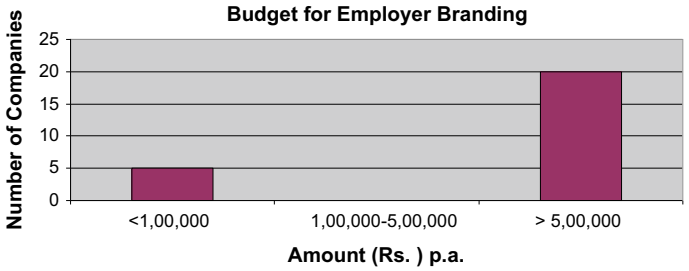


Fig. 4 The budget for employer brand
(*Source*: Author's representation)

budget. This shows that employer brand has started getting recognition as a separate head. Employer brand has become a part of the company's effort to establish among the potential work force as well as among the existing employees. This is required to retain and motivate employees as well as to attract new employees (Fig. 5).

The above diagram shows the graphical representation of the question of 13 (no.2) which pertains to the employer brand image of the company from employee's view point. Most of the organizations, 60% organizations, feel that they are well known in the eyes of potential employees' view point, and they do not have difficulty in attracting in employees. 40% of the organization feels that their employer brand is known, "good," in the eyes of potential employees' view point (Fig. 6).

The above chart shows the factors which are responsible for attracting the employees to their organizations. All of the companies knew that career growth was of paramount importance for attracting employees. So the organizations should provide career growth opportunities for the potential employees. The next in importance was the brand name of the employer, so people or potential employees are getting attracted toward the company because the company has a good brand name in the eyes of the potential employees. Salary, past track record of the company was not that important as compared to other factors. Few had the opinion that good working

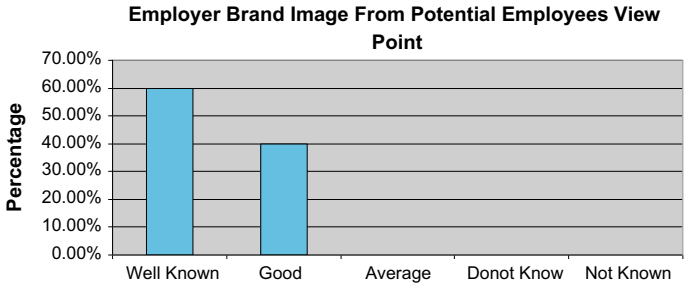


Fig. 5 Image of company from potential employees' view point
(*Source*: Author's representation)

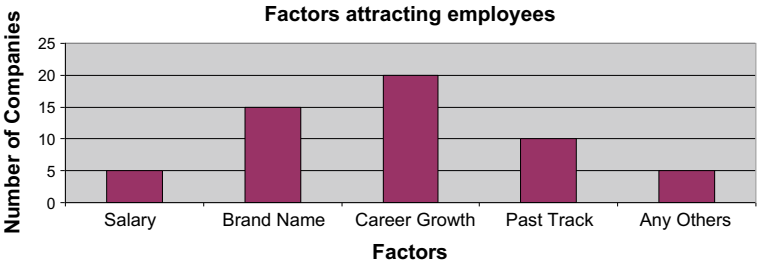


Fig. 6 Factors attracting employees
(*Source*: Author's representation)

environment and considering employees as human capital was also important ("What Is Employer Branding and How Can It Grow Your Business," 2018).

8 Conclusion

In today's time, there is a lot of competition for retention of talent, and those companies which have positioned themselves as a good employer or have a strong employer brand will be definitely in a strong position. These companies can even pick and choose from the best candidate. On the other side of the coin, having a strong employer brand will mean that the talent chooses them, so they need not to hunt for the talent.

The research on the employer brand has also revealed that employer branding is more relevant for IT companies in India, and most of the IT companies have strong employer brand in the market ("Employer Branding," 2019). This employer brands have helped these companies to attract the best talent in the market and to retain the same talent in the organization. In this research paper, we have highlighted some of the important points which are relevant for the IT companies as far as employer brand is concerned. The research has showed that IT companies have separate budget for employer brand, and they really do not have difficulty in attracting the **right talent** for their organization. These IT companies are well versed with the fact as to where the potential employees hang out, and what are the factors which the potential as well as current employees are looking in an organization.

With this, we conclude that employer brand is a concept which is not only known, but very well implemented concepts among Indian IT companies, and these companies have been able to differentiate themselves in the market place. This had led to the situation where they need not hunt for the talent, rather talent comes to them, making the recruitment process less costly and increasing the profitability for the organizations.

References

1. Velle J (2001) Employer branding: getting straight to the heart of the matter, working wire
2. Macioschek A, Katoen R (2007) Employer branding and talent-relationship-management: improving the organizational recruitment approach
3. Kotler P (2003). Marketing management (11 rev. edn)
4. Hodes global employer brand survey, June 2006
5. Corporate leadersip council (1999) The employment brand: building competitive advantage in the labor market
6. Melin E (2005) Employer branding: likeness and differences between external and internal employer brand images

A Short Paper on Web Mining Using Google Techniques

Kriti Gupta and Vishal Shrivastava

Abstract The quick improvement of data innovation throughout the most recent couple of years has brought about information development for a gigantic scope. Users make substance, for example, blog entries, tweets, informal community communications and photos; workers constantly make action logs; researchers make estimations information about the world we live in; and the web, a definitive store-house of information, has gotten dangerous concerning adaptability. To optimize web activities using different techniques can help in bringing the insights out for a clear big data picture.

Keywords Google Analytics · Web mining · Google trends · Graphical representation · Data mining

1 Introduction

The mining is extracting useful data or information from given insights. Web mining is the utilization of information mining strategies to discover data designs from web information. Web mining assists with improving the force of web Internet searchers by recognizing the pages and characterizing the web reports. Web mining is very useful to e-commerce websites and e-services.

Web mining expects to find and recover helpful and intriguing examples from large data sets, just as in the exemplary information mining. Huge information goes about as informational indexes on web mining. Web information incorporates data, reports, construction and profile. Web mining depends on two concepts characterized: process based and information driven. (in view of information regularly and commonly used) In the perspective on web mining information, web is utilized to remove information. As a rule, the use of web mining normally includes a few stages: gathering information, choosing the information before processing, information disclosure and examination.

K. Gupta (✉) · V. Shrivastava
Arya College of Engineering and IT Jaipur, Jaipur, India
e-mail: Kritigupta1497@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_62

Web mining is decomposed into following different subtasks :

1. **Searching resource**: finding web documents for retrieving information.
2. **Selection and preprocessing**: to select and automatically preprocess the required data retrieved from Internet resources.
3. **Generalization**: to identify similar patterns by checking all website at individual level and across altogether.
4. **Analysis**: to validate same or relevant patterns and identify likewise information for decision making [1].

It is vital to utilize information mining strategies to dissect information from the exercises did by guests to these sites. When all is said in done, online business and e-business have empowered online exchanges, and by creating enormous scope, ongoing information has never been simpler. Web content mining is the way toward extricating information from records and substance depiction. Web structure mining is the way toward getting information from association of the Internet along with the all other pages, etc. Mining is separated into three classes as follows:

- the web content
- the web structure, and
- the web utilization.

2 Web Mining Using Tools

Exactly when a visitor gets to any section of customer website (for instance www.abc.com), a requesting is made to the web specialist to show the page. The web laborer estimates the requesting and sends back the response to the program. By then, the program stacks the page. While stacking, it executes the JavaScript GA following code present in that site page. The standard reasoning of the JavaScript following code is accessible in the site visit work [2].

2.1 Web Mining Using Google Analytics

2.1.1 Architecture of GAT

1. Exactly when visitor gets to any portion of customer website, a sales is made to the web laborer to show the page. The web laborer estimates the sales and sends back the response to the program is showed up in the. Program stacks the page; while stacking, it thusly executes the JavaScript GA following code present in that site page.
2. The essential reasoning of the JavaScript GA following code is accessible in the track online visit() work.

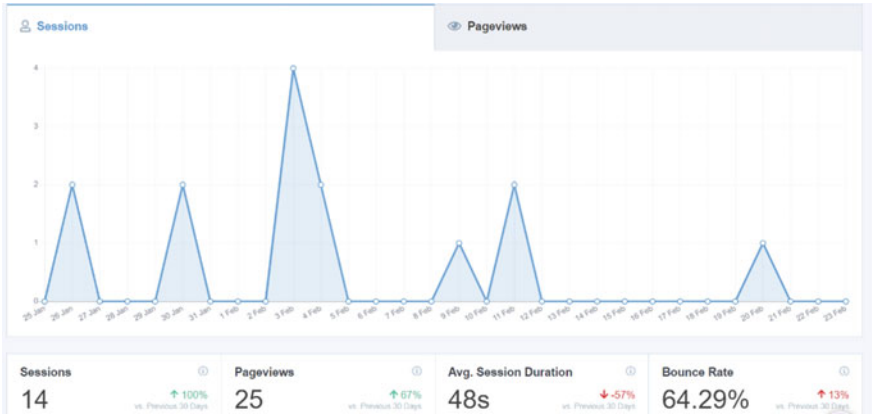


Fig. 1 Session analytics for the website

3. Google Analytics as of now eliminates customer treat regards. It invigorates/makes the treats as basic and sends the requesting to GAT specialist where the data is taken care of and after the planning of all the data then, at that point period of Google Analytics report.
4. This data is finally open for heads to find in customer's dashboard. This is to be certain straightforward cycle, and GA is one of the fundamental free examination instrument in the market open today. The data that Google Analytics uses to give all the information in customer reports comes from these sources:

    - The HTTP solicitation of the client
    - Browser/framework data
    - First-party treats [2] (Fig. 1).

The user can analyze the bounce rate and average session duration from here. The less the bounce, rate the more the visitors are coming and staying on your site. To keep the bounce rate low and increase duration, we can focus on creating more interlinking which will help user to stay and browse through the website (Fig. 2).

The above pie chart depicts the relation of how the traffic is coming and from where. The new and returning visitors are almost same, whereas the most visitors are coming from desktop. This suggests that blogging website can still have more visitors from web instead mobile. The design for website can be taken into consideration more to give better user experience in terms of loading and good speed (Fig. 3).

The above list suggests which is the most viewed page among all and is liked by people the most. Also the most viewed can be considered as trending to understand user interest.

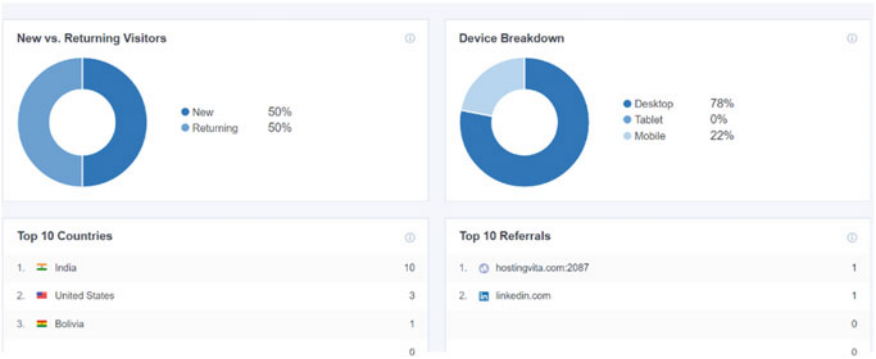


Fig. 2 Device-specific overview for the audience

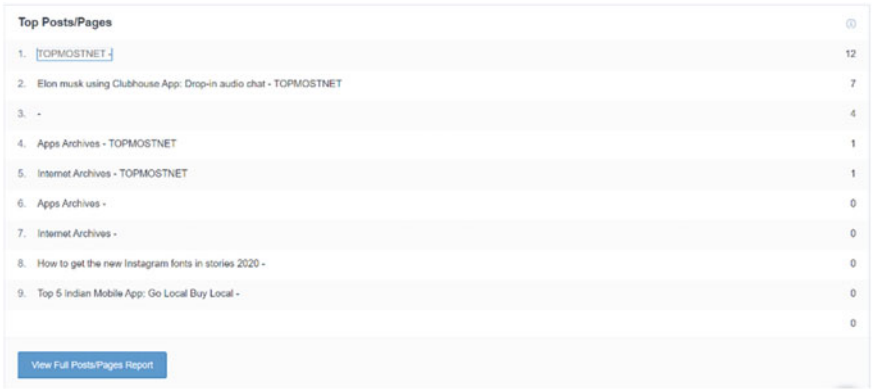


Fig. 3 Page-wise ranking as per the audience traffic on website

2.2 Web Mining Using Google Trends

Google patterns were not just worked to give a match to month to month catchphrase volume, and Google patterns were worked with a high-level degree of complexity, to create experiences which are visual and dynamic in nature. Google patterns are equipped for introducing a whole life pattern of a watchword expression, past, present and some sort of future as we may anticipate (Fig. 4).

The trends will depict that what keyword is when in demand and at what time. The keyword trending can be checked from here and people from which location are more likely to take interest in the same (Fig. 5).

The section will depict that which region people are more inclined to the searched keyword and what can be the next strategy and how to pitch accordingly (Fig. 6).

The above snapshot will tell the target audience range that up to what limit the queries are being searched. The queries can either go for specific number which

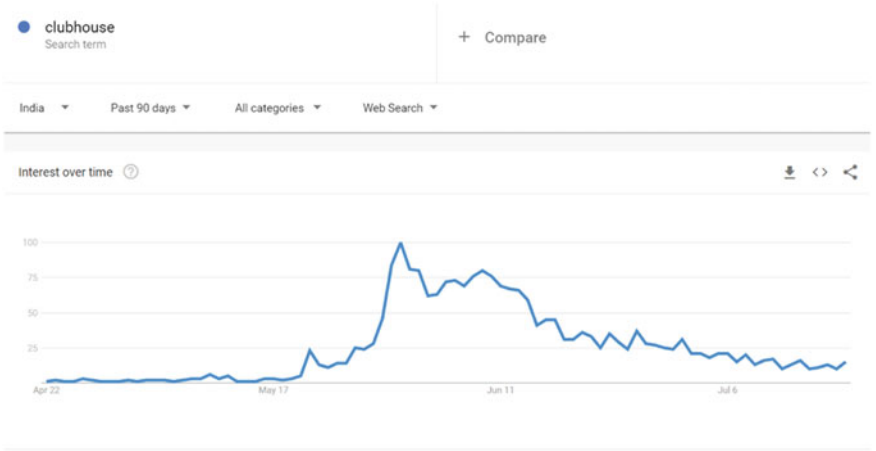


Fig. 4 Analyzing new trends

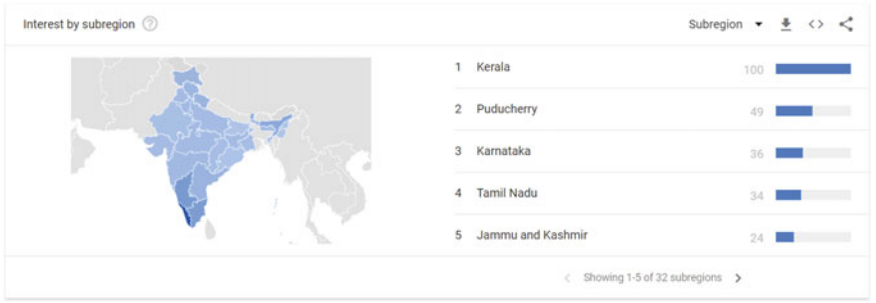


Fig. 5 State-wise demand

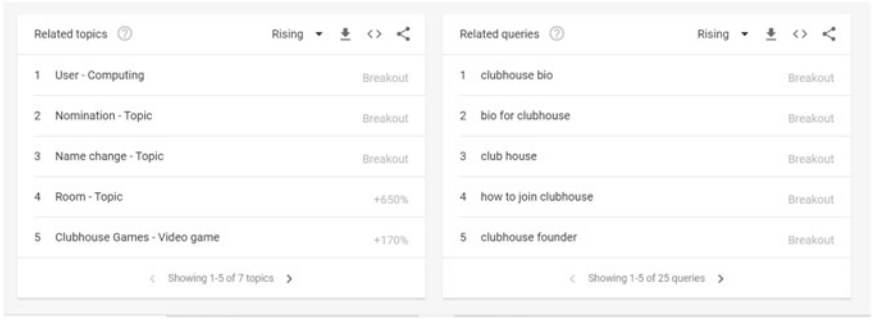


Fig. 6 Most searched items related to the keyword

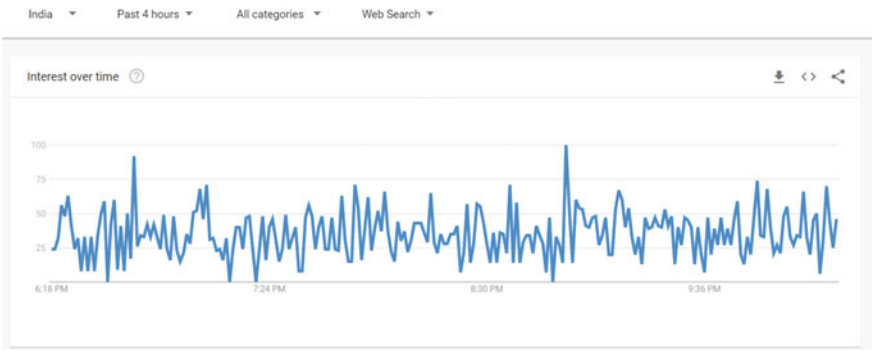


Fig. 7 Filter based on last searches

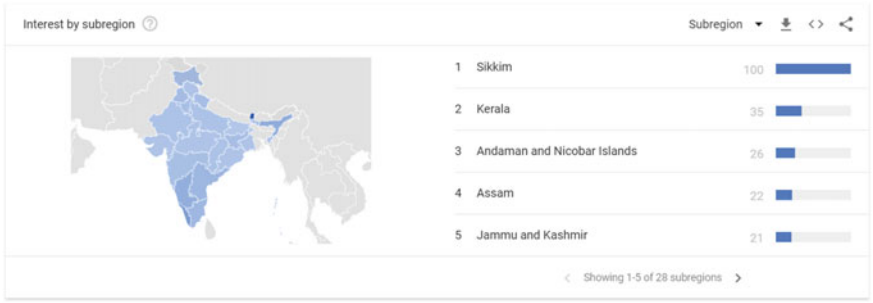


Fig. 8 Location based on last updated duration

means the search for that keyword is not infinite and is not good to focus, whereas the breakout result will suggest that these are searched n no. of times and can be a good fit for next article or blog (Figs. 7 and 8).

3 Conclusion

The checking of online questions can give understanding into human conduct, as this field is altogether and ceaselessly developing and will be demonstrated more than important in the future for evaluating social changes and giving ground to explore utilizing information that could not have been gotten to something else.

Google Trends is better on analyzing the activities taking place on the Internet, whereas calculating and estimations of it can be done using analytics.

Highlights of Google Analytics are as follows:

- It is liberated from cost.

- It gives tweaked reports as PDF, XML, CSV and TSV. In any case, it permits clients to have just one dashboard for each site. Along these lines, no numerous dashboards are permitted.
- As it is a free assistance, there is no assurance of utilization support when utilizing Google examination. Notwithstanding, Google 10 support oftentimes screens different discussions to help the clients.
- Users can just utilize one of the two of advantage/access (read just) levels for every site.
- Filtering is done from various perspectives including by space, IP address, site segment, referrer, page title and so on.
- Data is put away on Google worker.

Highlights of Google Trend are as follows:

- It is also free for use.
- Fast and free.
- Compare more number of terms all together.
- Can compare regions/country/location which will result good information and user behavior.

References

1. Uma Maheswari D, Marimuthu A (2013) A study of web usage mining applications and its future trends. Int J Eng Res Technol (IJERT) 2(9)
2. Thushara Y, Ramesh V (2016) A study of web mining application on E-commerce using google analytics tool. Int J Comput Appl 149(11):0975–8887
3. Ramya C, Shreedhara KS, Kavitha G (2011) Preprocessing: a prerequisite for discovering patterns in web usage mining process. In: 2011 international conference on communication and electronics information (ICCEI 2011), 978-1-4244-9481-1/11/$26.00 C 2011. IEEE
4. Aye TT (2011) Web log cleaning for mining of web usage patterns, 978-1-61284-840-2/11/2011. IEEE
5. bd Wahab MH, Mohd MN (2008) Data preprocessing on web server log for generalized association rule mining algorithm. World Acad Sci Eng Technol 48
6. Chitraa V, Davamani AS (2010) An efficient path completion technique for web log mining. In: IEEE international conference on computational intelligence and computing research 2010. ISBN: 97881-8371-362-7
7. Thakare SB, Gawali SZ (2010) A effective and complete preprocessing for web usage mining. Int J Comput Sci Eng (IJCSE) 02(03):848–851
8. Ivancsy R, Juhasz S (2007) Analysis of web user identification methods. World Acad Sci Eng Technol 34

Overview of Blockchain Technology and Its Application in Healthcare Sector

Neha A. Samsir and Arpit A. Jain

Abstract A blockchain is distributed ledger technology (DLT) which is used to store data securely. In the blockchain, data is stored as a peer-to-peer network. In this digital era, most of the business wanted to convert their data centers from centralized to decentralized. Blockchain can secure their data through a specific algorithm. The blockchain ensures data transparency, data security, and data integrity. Blockchain is known for cryptocurrencies, but it is beyond cryptocurrency. This paper aims to overview the architecture of blockchain, features of blockchain, and application of blockchain in the healthcare sector. The main goal of applying blockchain in the healthcare sector is for security and to achieve interoperability. Patient data is very sensitive and very crucial to provide security features. Due to COVID-19, most of the medical data is stored in electronic format, and storing their data in the blockchain is the best thing to make data secure and transparent.

Keywords Blockchain · Decentralization · Security · Challenges and risks · Health care

1 Introduction

Blockchain technology is a revolutionary technology for storing information. In the blockchain, once the information is stored, it will be difficult or impossible to change. It is also considered an immutable database. "Ledger" is one of the elements which is used to store information like a database. A blockchain contains a block that can hold encrypted transactions; each block can be chained to the other block in chronological order [1]. Blockchain can store a variety of transactions like data related to clinical, finance, and management. Blockchain concept can be derived from three different

N. A. Samsir (✉) · A. A. Jain
GLS University, Gujarat, India
e-mail: nehasamsir@gmail.com

A. A. Jain
e-mail: arpit.jain@glsuniversity.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_63

variants: decentralization, consensus mechanism, and cryptography [2]. Blockchain stores data in a unique way. Through decentralization, all nodes in the blockchain are responsible for the transaction. No centric authority can control the data, and there is no single entity that can be a reason for failure. All relevant data that can be stored on a ledger is also known as distributed ledger system.

Blockchain is the most innovative technology of the century. Blockchain is a continuously growing list of records and chronologically linked to each other. All blocks are stored or linked together in the chain of blocks, and the process of iterative all the blocks and validity of each transaction is called mining. The mining process is done through a specific algorithm. Each block contains a unique hash value [3].

Now, cryptocurrency is trending in the market which is created through blockchain technology. Bitcoin is the most famous cryptocurrency in the market. Bitcoin is a digital currency, and it will be managed by a consensus algorithm. Many problems arise in a physical transaction like a single point of failure and double-spent of money. To overcome the problem, digital cryptocurrency comes into the picture [4]. Blockchain maintains audit and traceability by linking the new block to the previous block of hash.

2 Literature Review

Data and information in the healthcare sector are very sensitive. Security is the key feature for the same. India faces many issues regarding the security of healthcare data. Few researchers had worked on the use of the blockchain technology in healthcare sector as follows:

Rim Ben Feikh and Mariam Lahami 2020 [5] proposed a basic concept of blockchain with its types and defining structure. The author states there are different types of use-cases where we can apply blockchain technology for storing medical records, remote patient monitoring, pharmaceutical supply chain, health insurance claims. Papers state some challenges in blockchain implementation which are scalability and implementation issues and latency issues. The author concludes to develop efficient applications and improve scalability and privacy issues.

Shahnaz et al. [2] proposed to implement healthcare sector data centralized to decentralize. There is some issue in centralized data that is interoperability, information asymmetry, and data breaches. Blockchain technology can overcome the issue than author describes the basics of blockchain technology and describing some challenges which are occurring in it.

Kitsantas et al. [1] proposed blockchain technology and different applications uses in business. Blockchain is working on smart contracts which are one type of computerized protocol which is useful to build trust between users and systems. Many applications are working based on smart contracts which are contract management, entertainment, health care, insurances, and blockchain Internet of Things (IoT). There are many other sectors that can adopt blockchain technology which is the blockchain for e-Government, blockchain in the financial industry, real-time accounting, and

business process management. There are many challenges and barriers which can occur during the implementation of blockchain.

Yaqoob et al. [6] proposed a basic concept of blockchain. According to Melaine Swan, blockchain is categorized into three phases which are 1.0, 2.0, and 3.0. Give a basic introduction of blockchain according to its features and ownership. Authors review some example who is using blockchain application in the healthcare sector which is MedRec, Gem, Healthbank, Guardtime. Paper states the major issue which is about healthcare stakeholders, which blockchain features are used to resolve the issue of record management, challenges.

Mikula and Jacobsen [7] proposed the application of blockchain in identity and access management using the hyperledger fabric framework. Generally, the healthcare sector uses Etherum, but in this paper, we aim to develop a healthcare data store through the blockchain with hyperledger fabric. Give generic information about the blockchain types and their structure and smart contracts. The authors created an application for the healthcare system. For the frontend, Angular JS is used, and for the back end, hyperledger fabric is used for creating a blockchain network.

3 Blockchain Technology

3.1 Key Features of Blockchain

Blockchain technology has many key features which are as follows.

- **Decentralization**: Blockchain is a decentralized network and open ledger. Ledger stores records of all the transactions, and every other in the network can go through with the transaction, as it is an open ledger system. All nodes in the network of systems can monitor systems that can access, store, and update the data [8]. Figure 1 shows the architecture centralized and decentralized network.

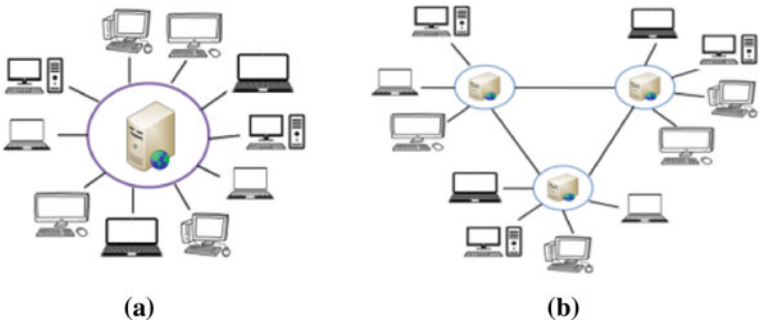


Fig. 1 **a** Centralized network, **b** Decentralized network

- **Anonymity**: A data transfer between one specific node to another node is anonymous which means no one knows the data. Then, it is more secure and reliable.
- **Consensus**: Every blockchain has a consensus to help the network to make decisions. The consensus algorithm is being made network trustless because of its algorithms. There are many consensus algorithms for blockchains over the globe.
- **Immutability**: Immutability means something cannot be changed or altered. The records in blockchain once stored are reserved forever and not easily modified.
- **Distributed Ledger**: Distributed ledger provides information about the transaction and users. It is responding very well to any suspicious activity or tamper. Every active node in the blockchain has maintained the ledger and participates in the validation.
- **Transparent**: In blockchain, all data is stored and a record is transparent which means it can prevent data from being altered or stolen. Through decentralization, every change on the blockchain is viewable and makes it more concrete.
- **Enhanced Security**: Decentralization and cryptography take another layer of protection to the user in the blockchain. All information in the blockchain is secure with a cryptographic algorithm.
- **Faster settlement**: Blockchain features make the transaction easier and faster. Past time, it takes lots of time for the transactions, but blockchain makes our life easier.

3.2 Block Structure

Blockchain is one type of sequence of records that holds the list of transactions or records. In a blockchain, all blockchain connected in an iterative manner and back to the previous block is known as the genesis block. The block consists of two parts block header and block body. Figure 2 shows the basic structure of the block where block header and transactions are mapped into the block body. Block $i + 2$ is having the hash value for the previous block $i + 1$. Block $i + 1$ and similarly the hash value of i. Through with they have connected. Table 1 represents the details of header parts.

3.3 Types of Blockchain

There are mainly three types of blockchain available: public (permission less), private (permissioned), and consortium (public, permission) [7]. In a public blockchain, no one individual can control the transactions recorded in the blockchain. This blockchain is highly resistant than easily anyone can join the network. All transactions in public blockchain are anonymous and transparent. Public blockchain are all tokens associated with them and designed to incentivize and participate in the network. Bitcoin is one type of public blockchain. In a private blockchain, all the

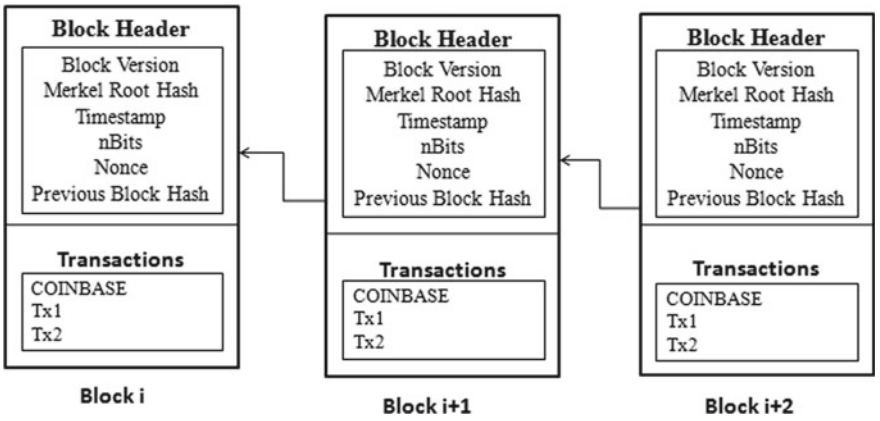


Fig. 2 Structure of blocks in blockchain

Table 1 Description of header parts

Field	Description	Size
Block version	A block version stores block version number to track particular software and protocol upgrade	4 bytes
Merkel root hash	It stores the hash of the root of the Merkel tree of this block's transactions	32 bytes (256 bits)
Timestamp	It stores the approximate creation time of this block	4 bytes
nBits	It targets threshold of a valid block hash	4 bytes
Nonce	It is counter used to consensus algorithm which usually starts with 0 and increases with every hash calculations	4 bytes
Previous block hash	It stores the reference to the hash value of the previous (parent) block in blockchain	32 bytes (256 bits)

transactions were secret which means data is not available publicly, but members are known, then participants need consent to join the network. Without permission, the participant did not write and read in the blockchain. Without permission, the participant did not write and read in the blockchain. Private blockchain is centralized compared to a public blockchain.

In an enterprise, a private blockchain is valuable who do not want to share their sensitive data, and it is used by large companies. Generally, it is used for supply chain and assets ownership. Ripple and hyperledger are an example of private blockchain. Consortium blockchain is managed or governed by a group rather than a single entity. This approach is all same benefits as private blockchain and is considered as a subcategory of private blockchain. Banking, supply chain uses the consortium blockchain. Corda is an example of consortium blockchain [3]. Table 2 represents properties of public, private, and consortium blockchain.

Table 2 Types of blockchain

Property	Public blockchain	Private blockchain	Consortium blockchain
Efficiency	Low	High	High
Consensus mechanism	All miners	One organization	Selected set of nodes
Centralized	No	Permissioned	Partial
Consensus process	Permission less	Permissioned	Permissioned
Read permission	Public	Public or restricted	Public or restricted
Transparency level	High	Low	Low
Scalability	Low	Medium	High
Immutability	Nearly impossible	Could be tampered	Could be tampered
Use-cases	Document validation cryptocurrency	Supply chain, asset ownership	Banking research supply chain

4 Applications of Blockchain

Many industries used blockchain technology for data protection. Blockchain improves data security, performance. Initially, blockchain is only in the finance sector, but now it is used in many other sectors also. Melanie Swan categorized blockchain in three phases like 1.0, 2.0, and 3.0. Blockchain 1.0 is the first phase of blockchain which can use as a cryptocurrency like bitcoin. Blockchain 2.0 can use for tracking the ownership of properties like copyright, financial records, and smart contracts. The major use of Blockchain 3.0 is in the science, health care, and education sector and in their applications [4]. Many sectors are now adopting blockchain technology. The application of blockchain is categorized in Fig. 3.

4.1 Blockchain in Healthcare Sector

In 2016, the national coordination for health information technology (ONC) suggested the utilization of blockchain in the healthcare sector. The healthcare sector is having many possibilities for the implementation of blockchain; few of them are for the implementation of blockchain; few of them are for storing patient data, pharmaceutical supply chain, remote patient monitoring, and insurance claim. Different users in healthcare (patients, hospitals, doctors, payers, practitioners) need to access and organize the same data without modification. Blockchain provides data authentication. A blockchain-based mobile application like MedRec, Gem, Guardtime has also gained interest to maintain medical data management, drug management, clinical trials, and data security [5]. Many key advantages can make the healthcare system impactful, decentralized management, immutable audit trail,

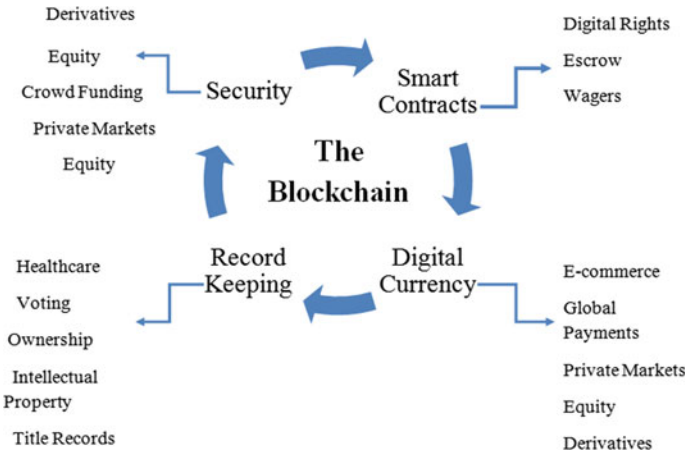


Fig. 3 Applications of blockchain technology

data provenance, robustness, availability, security, and privacy all are major benefits [6].

The study requires major three sections for the implementation of blockchain in health care [3]. Table 3 represents the three major components of the blockchain system and their challenges during implementation.

Table 3. Components of the system

Role	Advantages	Disadvantages	Challenges
People	(1) Secured view (2) Interoperability (3) Authorization	Data not changeable	No knowledge of new technology
Processes	(1) Tamper resistance (2) Transparency (3) Authentication (4) Efficiency (5) Immutable (6) Audit trail (7) Robust (8) Minimize human involvement	Access management	(1) Compliance (2) Maintain malicious traffic (3) Maintain accountability (4) Cost effectiveness
Technology	(1) Flexible (2) High performance and low latency (3) Authentication	(1) Cost (2) Complexity of consensus (3) Multi-dimensional data	(1) Cost effectiveness (2) Efficiency (3) Handling large amount of data (4) Manage 51% attack

5 Conclusion

Today, blockchain technology is trending in the market due to its security features. The researchers proposed an overview of blockchain technology which is defining blockchain architecture, defining key features of blockchain and different consensus algorithm. This study can develop a brief understanding of blockchain. The blockchain paradigm can provide robustness and decentralized management to the system. Many sectors have already adopted blockchain technology like education, health care, finance sector, etc. In this paper, we generally discussed blockchain technology and its application in the healthcare sector. The paper also elaborates challenges for the implementation of the blockchain system.

References

1. Kitsantas T, Vazakidis A, Chytis E (2019) A review of blockchain technology and its applications in the business environment. In: 2019, international conference on enterprise, systems, accounting, logistics & management. CRETE, Greece, China, 2019
2. Shahnaz A, Qamar U, Khalid A (2019) Using blockchain for electronic health records. In: 9th, IEEE Access, vol 7, pp 147782–147795, 19087857
3. Sadiku MNO, Eze KG, Musa SM (2018) Block chain technology in healthcare. PMCID—PMC6517629 vol 4, issue 5
4. Attaran M (2020) Blockchain technology in healthcare: challenges and opportunities. BSF-EHR: blockchain security framework for electronic health records of patients. Sensors 21(8):2865
5. Ben Fekih R, Lahami M (2020) Applications of blockchain technology in healthcare: a comprehensive study. Lecture Notes ComputSci 12157. ISBN : 978-3-030-51516-4
6. Yaqoob S, Khan MM, Talib R, Butt A, Saleem S, Arif F, Nadeem A (2019) Use of blockchain in healthcare: a systematic literature review. Int J Adv Comput Sci Appl (IJACSA) 10(5)
7. Mikula T, Jacobsen RH (2018) Identity and access management with blockchain in electronic healthcare records. In: 21st Euromicro conference on digital system design (DSD), pp 699–706. https://doi.org/10.1109/DSD.2018.00008
8. Zheng Z, Xie S, Dai H, Chen X, Wang H (2017) An overview of blockchain technology: architecture, consensus, and future trends. In: IEEE 6th international congress on big data, pp 557–564. ISBN: 978-1-5386-1996-4
9. Alla S, Soltanisehat L, Tatar U, Keskin O (2018) Technology in electronic healthcare system. In: Proceedings of the 2018 IISE annual conference, 2018
10. Kuo TT, Kim HE, Ohno-Machado L (2017) Blockchain distributed ledger technologies for biometrical and health care applications. J Am Med Inf Assoc 24(6):1211–1220
11. Suyambu GT, Anand M, Janakirani M (2019) Blockchain—a most disruptive technology on the spotlight of world engineering education paradigm. Procedia Comput Sci 172:152–157
12. Makridakis S, Christodoulou K (2019) Blockchain: current challenges and future prospects/applications

Digital Platforms and Techniques for Marketing in the Era of Information Technology

Roohi Sharma and Amala Siby

Abstract Digital marketing is the promotion of a product or service through at least one form of electronic media. This form of marketing is distinct from traditional marketing, but it uses some of the ideologies of traditional marketing. This research article examines the various technologies and platforms used in digital marketing that allow any organization or business to do this form of marketing and study what works for them and what does not. The article also explores the recent advancements in digital marketing due to the increase in users and the vast amount of data collected from these users. The two main advancements analyzed and discussed in this paper are machine learning (ML) and artificial intelligence (AI) tools.

Keywords Digital marketing · Digital tools · Digital platforms

1 Introduction

Traditional marketing refers to any marketing of goods or services that does not involve any digital aspect [1]. With many technological advancements, another form of marketing that came up in the 1990s is digital marketing, i.e., promotion of products and services over the internet or any other mode of electronic communication. It has a vital feature of interaction with customers at a much cheaper rate, which is why people consider digital marketing to be better than traditional marketing.

If utilized constructively, it can enable businesses to reach out to target customers without the hassle and generate brand awareness. Also, with inbound marketing ability, consumers find the product or service, rather than the other way around, like traditional marketing [2].

R. Sharma (✉) · A. Siby
CHRIST (Deemed to Be University), Bengaluru, India
e-mail: roohi.sharma@commerce.christuniversity.in

A. Siby
e-mail: amala.siby@christuniversity.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_64

2 Literature Review

Analyzing and understanding the articles and journal papers written by numerous well-established authors regarding digital marketing tools and techniques and how they can be implemented, this helps to understand the different perspectives and the research areas in the subject area of impact of digital marketing and how it can be used effectively by businesses as shown in Fig. 1.

2.1 Study of Digital Marketing

Businesses can benefit from this increase in internet users by using tools for digital marketing, ranging from search engine optimization (SEO) to social media marketing (SMM) that are becoming increasingly common in our advancing technology [3]. To summarize, digital marketing has an optimistic future with all its advantages and disadvantages for the long-term sustainability of the goods and services in the existing technological market [4]. However, this does not mean the end of traditional marketing as combining technology with many aspects of traditional marketing ideology will give more effective results [5].

2.2 Impact of Digital Marketing

An online marketing strategy could bridge the gaps between consumers to companies and virtual to physical. Such a strategy would allow for more efficient marketing approaches and a new relationship paradigm [6]. In general, digital marketing should be a crucial part of the approach for many consumers, as the information is readily available and best for reaching a wider audience and generating a global brand image [7].

Fig. 1 Literature review of various aspects of digital marketing

2.3 Future of Digital Marketing

The study examined various digital application areas that shape the digital marketing landscape and "marketing intelligence" from a marketing perspective [8]. Tools of ML and AI are being used extensively to collect and analyze data efficiently in marketing [9].

3 Research Gap

There have been a few articles about social media marketing as a technique. However, extensive research on other aspects of digital marketing which are not mainstream was neglected. Also, there is very little research in the area of applications of ML and AI in the digital market as it is a comparatively recent and upcoming topic.

4 Methodology

The research paper focuses on an extensive literature review through a systematic review process to understand the digital marketing tools, platforms and techniques used for business to better understand the past, present and future of digital marketing.

5 Objectives

The research objectives for this paper are:

- To identify the techniques and platforms used in digital marketing
- To find out how AI and ML are related to digital marketing
- To find out the impact and future of ML and AI on digital marketing.

6 Findings of Techniques and Platforms Used in Digital Marketing

In the current digital marketing landscape, a business can use numerous techniques and platforms to meet its marketing goals. The business will have to decide the techniques and platforms which are best suited for them as shown in Fig. 2.

Fig. 2 Techniques and platforms used in digital marketing

6.1 Social Media Marketing

Social media marketing enables a business to directly connect with its audience through various platforms that connect them to their audience like Facebook, LinkedIn, Instagram and Twitter. Examples of this technique can be employing social media influencers to promote their brand or actively posting and engaging with customers on their page [10].

6.2 Search Engine Optimization

SEO enables a business to improve its site by increasing its visibility for relevant searches. It gains the attention of internet users to a website by making sure that the website appears at top in a search engine like Google. This helps businesses in marketing their products and services as they are able to improve their brand awareness, and businesses can also use this technique on social media platforms by using relevant or trending hashtags.

6.3 Content Marketing

Content marketing consistently creates compelling content on platforms that interest and retain a clearly defined audience. A brand can use this content technique to

increase brand visibility by marketing itself to be in touch with its customers around the world. Like, Nike continuously running marketing campaigns for rising athletes [11].

6.4 Email Marketing

The process of a business sending an advertisement or promotional message of their products or services to a cluster of individuals using email is known as email marketing. An example of marketing through email is when companies send potential/existing clients, emails about upcoming offers to promote their brand.

6.5 Video Marketing

Video-based marketing works like content-based marketing. By creating quality commercials, a business can get people to purchase their products or services and ensure a positive and lasting brand image in the minds of its target audience.

6.6 Web Advertising

Web advertising or online advertising is another way businesses use the internet for advertising their products or services. Clickable advertisements that are put on popular webpages enable the businesses to market themselves more successfully as it will bring more footfall to the brands website and in turn get more sales.

6.7 Overall Personalization

Personalizing marketing is crucial for a business if they want to retain their customers. This technique involves using the customer's data, such as preferences, past purchases, to provide personalized advertisements.

6.8 Affiliate Advertising

Affiliate advertising is a technique where an affiliated and valuable spokesperson with an immense fan following promotes a brand in exchange for some commission.

This person can be a blogger, movie star, social media influencer, etc., who can influence his/her followers to buy the product and create brand awareness [12].

6.9 App Creation

App creation is a widespread digital marketing techniques and a crucial part of the marketing strategy for a business. It enables the customers of a business to download the app and then get notified of new products. It also gives them an accessible means of communication and purchase and makes the overall buying process easier.

7 Findings of ML and AI in Digital Marketing

In the present scenario of the marketing world, AI is already incorporated into many technologies such as chatbots and voice recognition, and there is no exemption that AI is swiftly becoming crucial and an integral part of the everyday digital world [13]. Here are some of the captivating applications of AI in marketing as seen in Fig. 3.

7.1 Focus on Consumer Behavior

AI is entirely focused on approaches that are data-driven for effective promotion of products and the related decision-making. Digital platforms collect and store a huge amount of data ranging from customer preferences to the amount of money spent as a

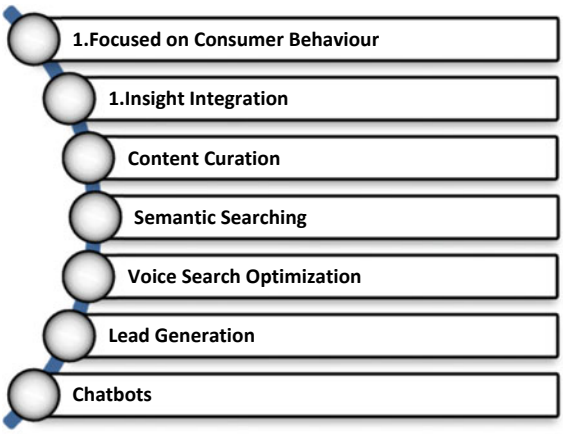


Fig. 3 Applications of artificial intelligence technology in marketing

part of analyzing the patterns of the customer which helps the business in developing customer profiles and assists the automated systems to target specific markets.

7.2 Insight Integration

AI is valuable to business because it can gather and integrate datasets from various software and other data gathering tools. The data can be used to find out and integrate the insights into consumer patterns. It will enable a business to target and customize digital ad campaigns based on buyer choices and purchasing habits.

7.3 Content Curation

AI can choose the most pertinent and tailored content for each unique visitor based on their likes and dislikes. A prime instance of this sort of AI technique is an e-commerce website such as Amazon "recommending" or showing similar products that the customer might like. AI does this by analyzing the buying behavior, finding patterns and using the patterns to recommend the products. Similar to suggestions on Netflix for TV series or movies.

7.4 Semantic Searching

Semantic searching refers to the ability of machines to provide customized outcomes. This is done through the machine's capability to understand user searches based on the context to offer a set of customized results. Unlike SEO, semantic searching shows the searches and tells why a person is searching for something.

7.5 Speech Recognition and Optimized Voice Search

The AI can handle a variety of search types, such as speech recognition. Voice search is one of the fastest-growing search methods. Businesses should incorporate a voice search strategy on their websites to better understand the customer [14].

7.6 Lead Generation

Lead generation is a bit like an automated recruiter. AI analyzes lots of data, builds on the information it already has and uses the program to find the ideal clients, customers, and colleagues. It can also forecast or rate how close it is to a particular lead.

7.7 Chatbots

Automated AI tools that are used for interacting with customers and buyers are known as chatbots. Chatbots are becoming increasingly common as they can be integrated on websites and used on other social media platforms [15].

7.8 Cybersecurity

The rise of disruptive digital technologies across industries and constant technological advancement in security solutions are critical drivers for the growth of AI in the cybersecurity market. Hence, it has become imperative for a business to focus on cybersecurity [16].

7.9 Automation and Personalization

ML and AI are being applied as a means of learning more about the customers, predicting buyer behavior and making decisions based on it [17]. The more it understands, the more businesses will be successful in targeting their marketing strategies toward what customers prefer [18].

8 Limitations

Since this is a qualitative study, it does not include the quantitative aspects such as the impact of digital marketing on buying behavior or company's cost. Also, the scope of the study is limited to application and future of AI and ML in digital marketing and does not consider the challenges or obstacles that can be faced while marketing in this digital era.

9 Conclusion

There have been immense advances in marketing, starting from traditional marketing methods. Today, AI is an incredible technology that can be very useful for a company. Using AI strategies for digital marketing will help companies deliver a better customer experience and be more effective in marketing. It will also ensure faster problem resolution for consumers. Companies must incorporate AI into their marketing strategies to understand their customers and gain a competitive advantage over their competitors.

References

1. Wolny J (2014) Marketing transformations: re-thinking marketing, digital first. J Direct Data Dig Market Pract 16:150–151
2. Jackson G, Ahuja V (2016) Dawn of the digital age and the evolution of the marketing mix. J Direct Data Dig Market Pract 17:170–186
3. Patil R, Patil MS, Bankar MV (2015) Digital marketing: the 21st century marketing tool. Int Res J Multidiscip Stud (IRJMS) 1(5):1–9
4. Desai DMV (2019) Digital marketing: a review. Int J Trend Sci Res Dev (IJTSRD) Special Issue:196–200
5. Gibson C (2018) The most effective digital marketing strategies & approaches: a review of literature. Int J Sci Res Publ (IJSRP) 8(2):12–16
6. Maria Teresa Pinheiro Melo Borges Tiago JMCV (2014) Digital marketing and social media: why Bother? Bus Horizons 57(6):703–708
7. Sathya P (2017) A Study on digital marketing and its impact. Int J Sci Res 6(2):866–868
8. Stone M (2015) Competitive marketing intelligence in a digital, data-based world. J Direct Data Digit Mark Pract 17:20–29
9. Lies J (2019) Marketing intelligence and big data: digital marketing techniques on their way to becoming social engineering techniques in marketing. Int J Interact Multim Artif Intell 5(5):134–144
10. Li J, Larimo J, Leonidou LC (2021) Social media marketing strategy: definition, conceptualization, taxonomy, validation, and future agenda. J Acad Market Sci 49:51–70
11. Belkhi RM (2018) The role of content marketing strategies in traffic generation: a conceptual model development. Digit Econ Emerg Technol Bus Innov 325:3–15
12. Jana Tabellion F-RE (2019) Influencer marketing and its impact on the advertised brand. In: Advances in advertising research X: multiple touchpoints in brand communication, pp 29-41
13. Theodoridis GDPK (2019) How artificial intelligence affects digital marketing. In: Springer proceedings in business and economics, no. Strategic innovative marketing and tourism, pp 1319–1327
14. Dellaert BGC, Shu SB et al (2020) Consumer decisions with artificially intelligent voice assistants. Market Lett 31:335–347
15. Adam M, Wessel M, Benlian A (2021) AI-based chatbots in customer service and their effects on user compliance. Electron Markets 31:427–445
16. Carley KM (2020) Social cybersecurity: an emerging science. Comput Math Organ Theory 26:365–381
17. Huang M-H, Rust RT (2021) A strategic framework for artificial intelligence in marketing. J Acad Market Sci 49:30–50
18. Davenport T, Guha A, Grewal D, Bressgott T (2020) How artificial intelligence will change the future of marketing. J Acad Market Sci 48:24–42

Tacit Knowledge Management in Engineering Industries: A Bibliometric Analysis

Pawankumar Saini and Pradnya Chitrao

Abstract Using a bibliometric methodology, this paper examines scientific research in the topic of tacit knowledge management in engineering companies. To that purpose, this report examines 705 publications from the Scopus database using several performance indicators, including total articles, total citations, and citations per paper. This paper also evaluates the most productive and well-cited authors, important subject areas, publication sources, countries, and institutions. The collection included publications published in scholarly journals between 1983 and 2021. The evolution of the research of tacit knowledge management is summarized in this publication. Our findings show that the most cited papers are from the United States and Norway, respectively. The most productive year in terms of published articles is 2010, whereas the most successful year in terms of citations is 2002. The findings could aid future research on this subject.

Keywords Tacit knowledge management · Knowledge management · Bibliometric analysis · Systematic review · Analysis of citations · Engineering industries

1 Introduction

Nonaka and Takeuchi talked about the volatile markets and the shifting competition, which is truer today than it ever was [1]. Updated production knowledge is crucial for firms to prevent diminishing returns on their stock of knowledge capital as technology evolves at a rapid pace [2]. When businesses recognize the value of information and

P. Saini (✉)
Symbiosis International University,Pune, India
e-mail: pawankumarsaini09@gmail.com

P. Chitrao
Symbiosis Institute of Management Studies,A Constituent of Symbiosis International University,Pune, India
e-mail: pradnyac@sims.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_65

begin to accumulate and preserve it, small- and medium-sized businesses can use it to investigate problems and develop solutions, resulting in a structure that facilitates efficiency and effectiveness [3]. A key part of knowledge management is knowledge sharing, which has become the leading factor for success in an organization [4]. A firm as a knowledge-creating organization is a crucial component of the knowledge-based view of the firm, which contends that knowledge and the ability to produce and use it are the most essential sources of a firm's long-term competitiveness [5]. Knowledge is a fluid combination of experience, values, contextual information, and expert insight that aids in the evaluation and assimilation of new information and events. This type of knowledge arises and is utilized exclusively in the minds of those who know [6]. Tacit knowledge (TK) has been identified as a crucial resource for long-term competitive advantage and company expansion [7]. Few companies make a concerted attempt to uncover TK, obscuring implicit knowledge [8]. Michael Polanyi talks about one of the key aspects of TK, that "we know more than we can tell" [9]. As per Haldin-Herrgard T, one of the main problems in TK sharing is the unconsciousness of TK in the minds of people and the difficulty of using it [10]. According to Howells, TK can be acquired in informal ways as it is non-codified and disembodied know-how [11]. As per Jones and Leonard, we need to find ways of converting TK to organization knowledge by considering the different moderating factors, so that many more acquire the knowledge and use it [12]. Paolino, Lizcano, Lopez, and Lloret say that TK is unlimited as it expands the more it is used [13]. As per Mohajan, a lot of TK creation and spread depends on individuals' social relationships and attributes [14].

We hope to gain a better understanding of the research that has been done so far and the areas in which it needs to be done further in the field of tacit knowledge management in the engineering sectors. We use bibliometric analysis to scan the arena of TK management in this research. Alan Pritchard used the term bibliometrics to describe the application of mathematical and statistical tools to books and other forms of communication [15, 16]. Academic journals, books, patents, proceedings, and other sources of information (i.e., citations, keywords, titles, journals, authors, institutions, and so on) provide an important sample for statistically performing science evaluation research through a methodical, transparent, and repeatable review process [15, 17–19]. It also aids in identifying the most prolific researchers and institutions, identifying alterations in discipline boundaries, and presenting the "big picture" of current research time [20, 21]. Not just for research, but also for policymaking and practice, the definition of intellectual structure and the research-front of scientific domains is critical [22]. It is becoming more popular as a technique for evaluating scholarly quality and output [23]. Journal-specific studies have also emerged, other than the topic-centric bibliometric analysis [24–29]. The paper's main contribution is a complete bibliometric analysis of "Tacit Knowledge Management in Engineering Industries" using the Scopus database, as well as the expansion of research in this field through time. Most significant writers, extremely productive authors, countries, most referenced disciplines, and highest contribution institutions have all been discussed.

2 Data Collection and Method

We collected the data from SCOPUS database, which is a widely referred repository [30]. The query for extracting the documents included the words "tacit", "knowledge", "engineer", and "industry", all with an Asterix so that all the related words like plural forms can be included. Further, we also included the words "firm" and "organization" as it is closely related to the word "industry". The query string thus generated on 3rd of July 2021 was: (TITLE (tacit*) AND TITLE (knowledge*) AND TITLE (engineer*) AND TITLE (industr*) OR TITLE (firm*) OR TITLE (organization*)). We obtained various tags from the Scopus database. The numerous performance characteristics collected from the bibliometric study, such as total articles, co-authors per article, articles per author, authors per article, total citations, and citations per article, are shown in this work.

3 Results and Discussion

The main information that emerged from the extracted documents based on the query was that articles made up 39% of the documents, conference papers 51%, and the remaining 10% consisted of books, book chapters, conference review, editorials, and reviews. The entire time span covered by Scopus was considered, which is 1983–2021. The 705 documents are available from 465 sources. The average citations per document is 17, and there are 21572 references made in these documents. A total of 1602 authors contributed toward research in this area. There are 166 documents written by single authors, which means that there are 539 co-authored documents. Multiple-author articles have a bigger impact than single-author articles since they are more likely to receive citations [31]. The collaboration index is 2.71. We have 1745 author's keywords, and there are 3868 keywords plus. In this section, we have conducted the performance analysis, which includes descriptive bibliometrics and evaluative bibliometrics [32]. We can get the research field's trends, influential authors, and most cited references, based on this [33]. We have presented bibliometric results of different performance parameters like research growth, most productive and highly cited authors, co-citation, keyword analysis, most sought after discipline, leading journals, country-wise analysis, institution-wise analysis, trending fields, and particularly influential papers.

3.1 Research Growth

Table 1 displays the total number of papers published (TP) and the total number of citations (TC) received by the articles throughout the course of the year. The number of papers published and the number of patents issued are essential indices

Table 1 Annual number of papers published and citations received

Year	TP	TC	Year	TP	TC
1983–1999	21	1174	2011	36	189
2000	10	550	2012	46	403
2001	11	677	2013	34	263
2002	16	1909	2014	45	177
2003	18	793	2015	36	226
2004	24	947	2016	37	375
2005	29	1329	2017	28	135
2006	41	796	2018	30	100
2007	34	356	2019	46	79
2008	39	774	2020	27	52
2009	31	355	2021	14	7
2010	52	320			

of scientific and technological production [34]. The very first paper recorded in this body of knowledge in the Scopus database was in 1983 by Glymour C., Stalker D.. TK management has received larger attention in the years starting from the year 1998, and it kept growing till 2006 after which the growth has been stable. This indicates a consistent interest in TK in industry. This could be because the number of researchers around the world is increasing, and due to internet, it has become convenient to carry out research [35–37]. The year 2010 saw the highest number of documents produced, totaling 52. From 1983 to 1999, the average number of papers per year was 1.23, whereas from 2000 to 2020, the average was 32 documents per year. We discovered that annual paper production is at an all-time high in the last two decades. Eugene Garfield coined the term "citation classic" (also known as "classic article" or "literary classic") to describe widely cited publications in a scientific domain. We will look at some of the citation classics, sometimes known as the "golden bullion of science" [38]. This could aid us in discovering critical information about the evolution of a particular discipline, allowing us to comprehend the history, present, and future of its scientific significance [39].

The two most cited papers have been authored by Asheim BT with a citation count of more than 760 in each of the articles. Other papers with more than 200 citations have been authored by Zucker LG, Demarest M, Almeida P, Mascitelli R, Bjrnson FO, and Lissoni. We can see that on an average, there have been 17 citations per document with 1.232 citations per year per document. There are more than 100 citations in 3.5% of the documents, 4.7% documents have citations between 50–100, and 17.5% have between 10 and 50 citations. Most citations are for publications between 2000 and 2010, indicating the focus that this subject received during this period. Articles published in the recent five years do not appear to have reached their maximum number of citations, and access to the first studies is not usually available to everyone; thus, they do not have a high number of citations [35, 36].

Table 2 Source title with the highest number of articles

Source title	TP	TC	C/D
Lecture Notes in Computer Science	21	89	4.2
Proceedings of the European Conference on Knowledge Management ECKM	21	22	1.0
Advances In Intelligent Systems and Computing	14	6	0.4
IFIP Advances in Information and Communication Technology	9	13	1.4
International Journal of Technology Management	9	151	16.8
ACM International Conference Proceeding Series	7	4	0.6
Automation In Construction	7	397	56.7
Communications In Computer and Information Science	7	6	0.9
Journal Of Construction Engineering and Management	7	204	29.1
Proceedings Of the Annual Hawaii International Conference On System Sciences	7	31	4

3.2 Journals, Authors, Institutions, and Countries That Are at the Top of Their Fields

The top 10 journals in relation to the quantity of papers published and the quantity of citations obtained are listed in Tables 1 and 2, respectively. The "Lecture notes in computer science" and "Proceedings of the European Conference on Knowledge Management ECKM" are the two sources with the most articles. However, the most important source that has made significant contribution around TK management in engineering industries is "Research Policy", which has received the highest citations (TC = 1212). The other significant sources with high citation per document (C/D) are "Journal of Technology Transfer" (C/D = 326) and "Management Science" (C/D = 330). Although the sources "Lecture Notes in computer science" and "Proceedings of the European Conference on Knowledge Management ECKM" have large number of articles, they have a low citation per document (Table 3).

An important perspective can be obtained from the data on authors, universities, and countries of papers citing the topic under study [40, 41]. Table 4 shows the top 10 authors, institutions, and countries based on the total number of papers (TP) produced, while Table 5 shows the top 10 based on the total citations (TC). R in the table denotes the ranking.

Table 3 Source title with the highest number of citations

Source title	TP	TC	C/D
Research Policy	5	1212	242
Journal of Technology Transfer	3	979	326
Management Science	2	659	330
Long Range Planning	3	657	219
Organization Science	2	429	215
Automation in Construction	2	397	199
Information and Software Technology	3	391	130
International Journal of Project Management	3	323	108
Journal of Product Innovation Management	2	266	133
Strategic Management Journal	2	242	121

Table 4 Authors, institutions, and countries with the most documents in the top 10

R	Authors	TP	TC	Institution	TP	TC	Country	TP	TC
1	Lin Y C	11	256	National Taipei University Of Technology	10	180	USA	200	3447
2	Wang L	6	76	University Of California	7	796	China	141	682
3	Liu Z	5	9	Loughborough University	7	157	UK	110	1068
4	Crcel-Carrasco J	4	8	Massachusetts Institute Of Technology	7	80	Japan	70	181
5	Heredia A	4	12	National Taiwan University	6	152	Sweden	44	1000
6	Liu X	4	6	University Of Twente	6	29	Brazil	41	66
7	Daneva M	3	13	Chalmers University Of Technology	6	21	India	40	123
8	Dingsyr T	3	305	Zhejiang University	6	5	Finland	35	98
9	Hamza H S	3	2	Georgia State University	5	256	Germany	33	132
10	Ibrahim R	3	8	Ching Yun University	5	76	Canada	31	337

We can observe from these two tables that the authors, institutions, and countries with the most article creation do not necessarily have the most citations. National Taipei University of Technology, for example, published the most articles on TK management in engineering businesses. Despite being the leading article publisher in terms of quantity, i.e., the number of articles published, this institute has not been mentioned as frequently as other institutes. University of Oslo and University of Lund, on the other hand, have been cited in 1548 publications, whereas they published only two articles each. This shows how the number of citations received by a highly cited article affects the ranks of authors, universities, and countries. Lotka devised a criterion to classify author productivity: small producers publish one article, medium producers publish two to nine articles, and large producers write more than nine articles [42]. In this case, there is only one large producer, whereas there are 48 medium producers. When we look at the authors who have contributed the most, we see Lin YC publishing 11 articles, Wang L with 6 articles, Liu Z with 5 articles, and Crcel-carrasco J, Heredia A, and Liu X with 4 articles each. There are 8 other authors who have published 3 articles each. While Lin YC has received over 250 citations and Wang L has received over 70, the remaining authors have received less than 15 citations. We see that despite low productivity by Asheim BT, only two articles: "Knowledge bases and regional innovation systems: Comparing Nordic clusters" and "Regional innovation systems: The integration of local 'sticky' and global 'ubiquitous' knowledge", his total citation count is above 1500, which is the highest [43, 44]. The other two articles with the third and fourth highest citations are by Zucker, Darby, and Armstrong: "Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology" and by Demarest: "Understanding knowledge management" [45, 46].

3.3 Relationship Between Leading References by Leading Authors and the Top Key Words

Figure 1 shows an interesting three plot of highest cited references (CR), highly productive authors (AU), and the top keywords associated by Scopus (ID) which have been extracted. We see that article written by Nonaka and Takeuchi: "Knowledge creating company" [1], Nonaka I.: "A dynamic theory of organizational knowledge creation" [47], and Davenport and Prushak: "Working knowledge: How organizations manage what they know" [6] have been cited the most among the other citations. These authors have influenced the research in this body of knowledge. Lin Y.C. is seen as the prominent contributor among all the other authors. It is seen that Lin Y.C. has largely researched with the construction industry addressing knowledge management, project management, and knowledge-based system. He has referred the work of Davenport T.H. and Prushak I to a large extent. Dingsyr T. has largely contributed to the areas of knowledge management, TK, software engineering, and knowledge engineering. Liu Z has prominently explored the TK space.

Table 5 Authors, institutions, and countries with the most citations in the top 10

R	Authors	TP	TC	Institution	TP	TC	Country	TP	TC
1	Asheim B.T.	2	1548	University Of Oslo	2	1548	USA	200	3447
2	Zucker L.G.	2	586	University Of Lund	2	1548	Norway	24	1262
3	Demarest M.	1	380	University Of California	7	796	United Kingdom	110	1068
4	Bresnen M.	1	290	Norwegian University Of Science And Technology	4	382	Sweden	44	1000
5	Almeida P.	1	266	University Of Warwick	2	290	China	141	682
6	Mascitelli R.	2	241	Georgetown University	1	266	Canada	31	337
7	Bjørnson F.O.	1	227	Georgia State University	5	256	Italy	20	331
8	Lissoni F.	1	216	University Of Illinois At Urbana-Champaign	2	252	Switzerland	20	315
9	Boiral O.	1	200	Technology Perspectives	2	241	Spain	31	242
10	Hoetker G.	1	198	University Of Brescia	1	216	Australia	29	205

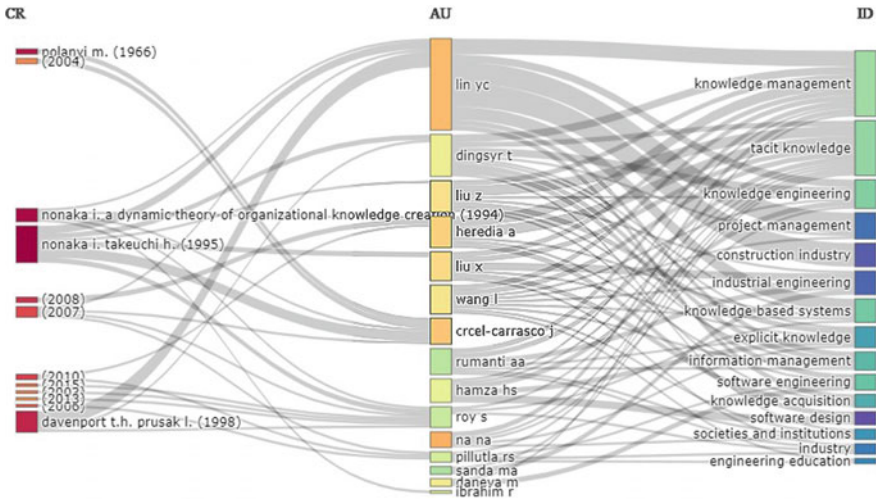


Fig. 1 Relationship between Leading cited references, leading authors, and the top keywords

Knowledge management, TK, and knowledge engineering are the keywords in most of the top producing authors. In the set of articles extracted from Scopus for this study, there were 3870 different keywords in the keywords plus. In the figure, we see the top keywords used by the authors. The most relevant words are knowledge management, TK, and knowledge engineering. The dominance of these words gives an idea that the subject area is dealing a lot in TK management and knowledge-based systems. The trending topics of a 20-year span starting from 2000 extracted from Scopus tell an interesting story. Knowledge management was trending weakly till about 2007 and then started picking up to become prominent by 2012. Topics like knowledge engineering, information management started trending around 2008. Other topics that further picked up were knowledge acquisition, knowledge-based systems, explicit knowledge, and knowledge economy. Knowledge transfer and TK were trending at their peak during 2014. The topics that have been trending in the recent few years are human resource management, safety engineering, knowledge extraction, maintenance, technological development, Industry 4.0, etc. This indicates the new areas in which TK management is being studied.

4 Conclusions, Limitations, and Future Research

We conducted a thorough bibliometric analysis in the topic of "Tacit Knowledge Management in Engineering Industries" in this research. The use of bibliometric analysis assisted in the discovery of trends and developments in this field. Lin Y.C. is the database's most productive author. In addition, Asheim B.T. continues to be the most influential author, with the most citations, followed by Zucker L.G. and Demarest M. The main subjects of discussion in this bibliometric analysis are knowledge management, TK, and knowledge engineering. The most common source of publications is "Lecture Notes in Computer Science", followed by "Proceedings of the European Conference on Knowledge Management". The most productive countries are the United States and Norway, with National Taipei University of Technology in Taiwan producing the most papers. The bibliometric study's weakness is that it is dependent on citations and the quantity of papers. The quantity of published work is represented by the number of articles and citations; however, citations do not indicate quality. Papers with a low number of citations may not necessarily have a low scientific merit.

This study can be quite valuable for researchers who are just getting started in this field and want to start with a literature review. The analysis drew attention to the regions that were being investigated. It is worth mentioning that the research has primarily concentrated on project management, the construction sector, and software engineering; nevertheless, TK is present in every organization, and it would be worthwhile to investigate functions and sections of the engineering industry that have yet to be investigated. Employee knowledge sharing (which is not well explored in this body of knowledge) can be critical for firms to focus on in order to discover and address problems quickly and efficiently, which leads to innovation [48, 49]. The use of information technology to capture or even to digitalize TK does not seem to figure

prominently in this body of research. It can be researched further so that the industry can benefit in making the not so visible TK more visible [50]. Based on the trending topics, future researchers could study the areas which are evolving, like Industry 4.0, maintenance, and knowledge extraction. Presently, web-based technologies are being widely used by many organizations for interactions [51]. Researchers can concentrate on the components of knowledge sharing that can be accomplished using online platforms, as well as the implications for TK sharing. We have just looked at the commonly used Scopus database for bibliometric analysis; however, there are additional databases worth looking at, such as open-access journals. Other indexing databases, such as Web of Science and Google Scholar, can be used to undertake further analysis.

Acknowledgements We would like to acknowledge Symbiosis International University for providing us the access to the library.

References

1. Nonaka I, Takeuchi H (1995) Knowledge-creating company. Knowledge-Creating Company, no Dec 1991
2. Woods JG (2019) From decreasing to increasing returns: the role of tacit knowledge capital in firm production and industrial growth. J Knowl Econ 10(4):1482–1496. https://doi.org/10.1007/s13132-016-0351-2
3. Gholami MH, Asli MN, Nazari-Shirkouhi S, Noruzy A, Investigating the influence of knowledge management practices on organizational performance: an empirical study
4. Fauzi MA, Paiman N (2020) A critical review of knowledge sharing in various industries and organizations. Int J Sci Technol Res 8. Available: www.ijstr.org
5. Nonaka I, Toyama R, Nagata A (2000) A firm as a knowledge-creating entity: a new perspective on the theory of the firm. Ind Corp Chang 9(1):1–20
6. Davenport TH, Prusak L (1998) Working knowledge: how organizations manage what they know. Harvard Business Press
7. Stenmark D (2000) Enhancing organizations. The quest for fast and effective value creation. Intellectual Bandwidth
8. Boiral O (2002) Tacit knowledge and environmental management. Available: www.lrpjournal.com
9. Polanyi M (1966) The tacit dimension, knowledge in organizations. Prusak L (ed)
10. Haldin-Herrgard T (2000) Difficulties in diffusion of tacit knowledge in organizations. J Intell Capital 1(4):357–365. https://doi.org/10.1108/14691930010359252
11. Howells J (1996) Tacit knowledge, innovation and technology transfer. Technol Anal Strategic Manage 8(2):91–106. https://doi.org/10.1080/09537329608524237
12. Jones K, Leonard LNK (2009) From tacit knowledge to organizational knowledge for successful KM, pp 27–39. https://doi.org/10.1007/978-1-4419-0011-1_3
13. Paolino L, Lizcano D, López G, Lloret J (2019) A multiagent system prototype of a tacit knowledge management model to reduce labor incident resolution times. Appl Sci (Switzerland) 9(24):2019. https://doi.org/10.3390/app9245448
14. Mohajan HK (2016) Sharing of tacit knowledge in organizations: a review. Available: http://www.openscienceonline.com/journal/ajcse
15. Pritchard A (1969) Statistical bibliography or bibliometrics? J Docum 25(4)
16. Pourkhani A, Abdipour K, Baher B, Moslehpour M (2019) The impact of social media in business growth and performance: a scientometrics analysis. Int J Data Netw Sci 3(3):223–244

17. Andrés A (2009) Measuring academic research: how to undertake a bibliometric study. Elsevier
18. Diodato VP (1994) Dictionary of bibliometrics psychology press. The Haworth Press, Binghamton
19. Gutiérrez-Salcedo M, Martínez MÁ, Moral-Munoz JA, Herrera-Viedma E, Cobo MJ (2018) Some bibliometric procedures for analyzing and evaluating research fields. Appl Intell 48(5):1275–1287. https://doi.org/10.1007/s10489-017-1105-y
20. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F (2011) An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the fuzzy sets theory field. J Informet 5(1):146–166. https://doi.org/10.1016/j.joi.2010.10.002
21. Crane D (1977) Social structure in a group of scientists: a test of the 'invisible college' hypothesis. In: Social networks, Elsevier, pp 161–178
22. Aria M, Cuccurullo C (2017) bibliometrix: an R-tool for comprehensive science mapping analysis. J Informet 11(4):959–975. https://doi.org/10.1016/j.joi.2017.08.007
23. Amsterdam V, Ioad6 A (1995) New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications
24. Yu D, Xu Z, Pedrycz W, Wang W (2017) Information sciences 1968–2016: a retrospective analysis with text mining and bibliometric. Inf Sci 418:619–634
25. Yu D, Xu Z, Kao Y, Lin C-T (2017) The structure and citation landscape of IEEE Transactions on Fuzzy Systems (1994–2015). IEEE Trans Fuzzy Syst 26(2):430–442
26. Muhuri PK, Shukla AK, Janmaijaya M, Basu A (2018) Applied soft computing: a bibliometric analysis of the publications and citations during (2004–2016). Appl Soft Comput 69:381–392
27. Laengle S et al (2017) Forty years of the European Journal of Operational Research: a bibliometric overview. Eur J Oper Res 262(3):803–816
28. Janmaijaya M, Shukla AK, Abraham A, Muhuri PK (2018) A scientometric study of neurocomputing publications (1992–2018): an aerial overview of intrinsic structure. Publications 6(3):32
29. Cobo MJ, Martínez M-Ó, Gutiérrez-Salcedo M, Fujita H, Herrera-Viedma E (2015) 25 years at knowledge-based systems: a bibliometric analysis. Knowl-Based Syst 80:3–13
30. Albort-Morant G, Ribeiro-Soriano D (2016) A bibliometric analysis of international impact of business incubators. J Bus Res 69(5):1775–1779
31. de Granda-Orive JI et al (2009) Redes de colaboración científica internacional en tabaquismo: análisis de coautorías mediante el Science Citation Index durante el periodo 1999-2003. Gaceta Sanitaria 23(3):222e34–222e43
32. Zemigala M (2019) Tendencies in research on sustainable development in management sciences. J Clean Prod 218:796–809
33. Bensalem A, Kin V (2019) A bibliometric analysis of reverse logistics from 1992 to 2017. Supply Chain For: Int J 20(1):15–28
34. Narin F, Hamilton KS (1996) Bibliometric performance measures. Scientometrics 36(3):293–310. https://doi.org/10.1007/BF02129596
35. Merigó JM, Mas-Tur A, Roig-Tierno N, Ribeiro-Soriano D (2015) A bibliometric overview of the Journal of Business Research between 1973 and 2014. J Bus Res 68(12):2645–2653
36. Merigó JM, Gil-Lafuente AM, Yager RR (2015) An overview of fuzzy research with bibliometric indicators. Appl Soft Comput 27:420–433
37. Merigó JM, Muller C, Modak NM, Laengle S (2019) Research in production and operations management: a university-based bibliometric analysis. Glob J Flex Syst Manag 20(1):1–29
38. Smith DR, Leggat PA (2008) Ten citation classics from the Australian and New Zealand Journal of Public Health. Australian New Zealand J Public Health 32(2):105–106. https://doi.org/10.1111/j.1753-6405.2008.00183.x
39. Martínez MA, Herrera M, López-Gijón J, Herrera-Viedma E (2014) H-Classics: characterizing the concept of citation classics through H-index. Scientometrics 98(3):1971–1983. https://doi.org/10.1007/s11192-013-1155-9
40. Amirbagheri K, Núez-Carballosa A, Guitart-Tarrés L, Merigó JM (2019) Research on green supply chain: a bibliometric analysis. Clean Technol Environ Policy 21(1):3–22. https://doi.org/10.1007/s10098-018-1624-1

41. Esfahani HJ, Tavasoli K, Jabbarzadeh A (2019) Big data and social media: a scientometrics analysis. Int J Data Netw Sci 3(3):145–164. https://doi.org/10.5267/j.ijdns.2019.2.007
42. Lotka AJ (1926) The frequency distribution of scientific productivity. J Wash Acad Sci 16(12):317–323
43. Asheim BT, Coenen L (2005) Knowledge bases and regional innovation systems: comparing Nordic clusters. Res Policy 34(8):1173–1190. https://doi.org/10.1016/j.respol.2005.03.013
44. Asheim BT, Isaksen A (2002) Regional innovation systems: the integration of local 'sticky' and global 'ubiquitous' knowledge. J Technol Transf 27(1):77–86
45. Zucker LG, Darby MR (2001) Capturing technological opportunity via Japan's star scientists: evidence from Japanese firms' biotech patents and products. J Technol Transf 26(1):37–58
46. Demarest M (1997) Understanding knowledge management. Long Range Plan 30(3):374–384
47. Nonaka I (1994) A dynamic theory of organizational knowledge creation. Organ Sci 5(1):14–37
48. Chopra M, Gupta V (2019) Linking knowledge management practices to organizational performance using the balanced scorecard approach. Kybernetes
49. Thomas A, Gupta V (2021) Tacit knowledge in organizations: bibliometrics and a framework-based systematic review of antecedents, outcomes, theories, methods and future directions. J Knowl Manage
50. Thomas A, Chopra M (2020) On how big data revolutionizes knowledge management. In: Digital transformation in business and society. Springer, pp 39–60
51. Zutshi A, Mendy J, Sharma GD, Thomas A, Sarker T (2021) From challenges to creativity: enhancing SMEs' resilience in the context of COVID-19. Sustainability 13(12):6542

Image Processing Techniques in Plant Disease Diagnosis: Application Trend in Agriculture

Debangshu Chakraborty and Indrajit Ghosh

Abstract In agriculture, plant diseases cause significant economic losses every year and are major threats against food security. To control the infestation at the earliest stage and minimize crop loss, precise diagnosis of plant diseases is the only way. The diagnostic process usually relies on the in-field visual identification by agricultural experts with rich experience. However, for a large number of underdeveloped and developing countries, human experts are scarce and expensive. As an alternative, image processing techniques are becoming popular for automated plant disease diagnosis in agriculture. Several image processing systems have been proposed worldwide during the last one and a half decades. Aside from a mere literature survey, this paper presents a statistical study of the application trend of various image processing techniques used to design image processing systems for plant disease diagnosis in agriculture. This study will be very beneficial for aspirant researchers in the future to design such systems.

Keywords Image processing · Plant disease diagnosis · Application trend · Segmentation techniques · Feature extraction techniques · Machine learning techniques

1 Introduction

The diseases of the plants play an antagonistic role in agriculture. Plants are often infested by several diseases, which reduce both the quantity and quality of agricultural products. Plant diseases cause significant economic losses to farmers every year throughout the world and are major threats against food security. Proper diagnosis of diseases at the earliest stage is very essential to take appropriate control measures.

D. Chakraborty · I. Ghosh (✉)
Agro-Computing Research Laboratory, Department of Computer Science, Ananda Chandra College, Jalpaiguri 735101, India
e-mail: ighosh2002@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_66

Most of these diseases cause site-specific signs and symptoms to a particular species of plant. The diagnostic process usually relies on the in-field visual identification based on signs and symptoms observed by agricultural experts with rich experience. However, for the developing countries having widespread cultivation, the human experts are scarce and expensive.

As graceful alternatives, several image-based intelligent systems have been proposed worldwide during the last one and a half decades [1]. Automated systems using machine vision hybridized with machine learning techniques are found to be very prosperous in plant disease diagnosis.

In such typical systems, the digital images of an infested part of a plant are acquired and various image processing techniques such as color space conversion, de-noising, sharpening/blurring are performed to enhance the quality of the image. A segmentation technique is applied to segregate the region of interest from the image. In the next step, the significant symptomatic features are extracted. A machine learning component uses the extracted features as input and diagnoses (classify) the concerned disease(s) accordingly. Various classifiers have been used for this purpose with their capabilities and limitations.

This paper projects the application trend of such techniques (other than deep learning) in plant disease diagnosis proposed over the last one and a half decades throughout the world. Hopefully, it will be helpful to future researchers to find out new avenues of designing such systems with better efficiency.

The rest of this paper is organized into five sections. Section 2 describes the methodology used to perform the study. An extensive survey has been made on the application of image processing techniques designed for plant disease diagnosis in agriculture over the last one and a half decades. These observations are analyzed to mine the application trends of these techniques, which is presented in Sect. 3. Discussions are provided in Sects. 4, and 5 contains the conclusions.

2 Methodology

This paper presents the study of the application trend of various intelligent techniques to design various image-based plant disease diagnosis systems in agriculture. Two principal factors were considered in the present study; quality and accessibility of the publications and reputation of the publication agencies. A total of 250 research papers have been accessed from reputed publishing houses such as Springer, ACM, Elsevier, Wiley, IEEE, and Taylor Francis. Google Scholar, SCOPUS, Microsoft Academic, Crossref were utilized as the major indexing sources. From these 250 published papers, 99 papers have been identified and considered relevant to the present study covering the last 16 years (2005–2020). Finally, selected 99 publications are grouped in four sets considering the year of publication within four-year time span (2005–2008, 2009–2012, 2013–2016, and 2017–2020). The outcome of the study is presented in graphs where the horizontal axis represents the time scale

Table 1 Segmentation techniques

Techniques	Description
Similarity based	In similarity-based segmentation, the image is partitioned by using several techniques like region growing, thresholding, region splitting, and merging [2–23]
Discontinuity based	Discontinuity-based segmentation can be done based on discontinuity in the pixel intensities of an image. In discontinuity-based segmentation, the image is partitioned based on sharp and abrupt discontinuities in the pixel intensity like edges [18, 24–28]
K-means clustering based	K-means clustering divides the image into a set of regions and one or more of those regions will have majority of diseased parts of the image. This algorithm clusters the pixels into k-classes using a feature set. This clustering is done by minimizing the sum of distance squared between the data object and clusters [1, 29, 29–52, 52–73]
Other techniques	Apart from these techniques, some other techniques were also used such as Bayes classifier, self-organizing-map, back propagation neural network, k-NN, fermi energy-based, and watershed segmentation techniques [18, 49, 74–84]

of 16 years and the vertical axis denotes the total number of applications of the respective technique.

3 Application Trend Study

3.1 Application Trend of Segmentation Techniques

Segmentation of complex images is one of the most challenging tasks in image processing. Segmentation accuracy determines the success or failure of the system. Various segmentation techniques are presented in Table 1, and the graphical representation of the application trend is presented in Fig. 1.

3.2 Application Trend of Feature Extraction Techniques

Feature extraction techniques based on texture features. Texture exhibits spatial structure or distribution of image intensity. There are three approaches to extract texture features: statistical, structural, and spectral. A very short description of these approaches is provided in Table 2, and the application trend is presented in Fig. 2.

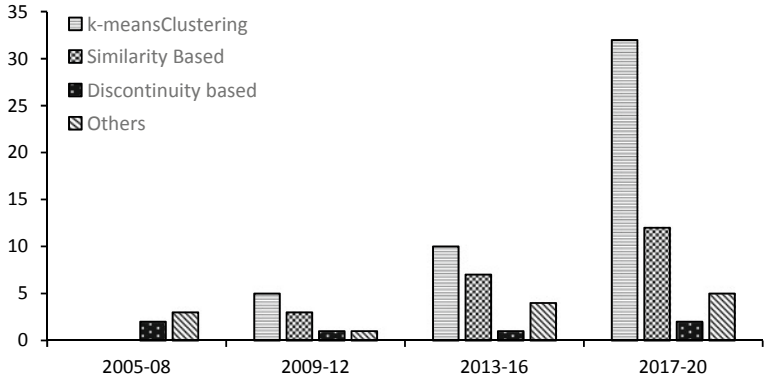


Fig. 1 Segmentation techniques

Feature extraction techniques based on color features. There are several techniques that are used to extract color features from the image of the disease-infested area of a plant, as described in Table 3. The application trends of all color feature extraction techniques are presented in Fig. 3.

Feature extraction techniques based on shape features. The morphological observations of the disease-infested area of a plant are very informative regarding the identification of diseases. Shape feature extraction techniques extract the relevant information from these morphological observations. Various shape feature extraction techniques are described in Table 4, and their application trends are presented in Fig. 4.

3.3 Application Trend of Machine Learning Techniques

For plant disease identification, various classifiers have been used in last one and a half decade with their own capabilities and limitations. The classifiers which are significant and widely used in image-based plant disease identification systems are described in Table 5. The application trends are presented in Fig. 5.

4 Discussion

This study has manifested a clear picture of the application trend of various techniques applied for image processing in agricultural plant disease diagnosis during the past one and a half decades. It is revealed that for segmentation, the k-means clustering technique is the most popular one with a gradual increase in the application. It is observed from Fig. 3 that among all the color features, the extraction techniques based

Table 2 Feature extraction techniques based on texture features

Techniques	Description
Statistical approach	This approach includes matrix-based methods like GLCM, GGCM, GLRM, CDTM, auto-correlation function, texture spectrum statistics, and semi variance graph methods [2, 3, 6, 9, 10, 12–14, 16, 17, 19, 21, 22, 26, 28, 29, 34, 36, 43–45, 47–50, 52–59, 61–63, 65–70, 72–74, 80–83, 85–93]
Based on structural features	These techniques represent texture as a collection of well-defined texture primitive and the structure and spatial relationships among the primitives. There are two approaches: space field-based and filter-based. Space field-based techniques are cluster analysis, pixel analysis, and morphological methods. Filter-based techniques includes quadrature mirror filtering, dyadic Gabor filtering, and laws texture extraction methods [77, 80, 92]
Based on spectral features	In this approach, first, the image is transformed to frequency domain and then analyzed. These techniques are wavelet transform method, Gabor transform, Ridget transform, curvelet transform, random transform, discrete cosine transform, local Fourier transform, and local Walsh transform [3, 20, 47, 49, 62, 89, 92, 94]
Other techniques	Other texture feature extraction techniques include local binary patterns, complete local binary patterns, chromaticity moments, Lacunarity, and calculating fractal dimension using box counting [2, 33, 37, 42, 76, 95, 96]

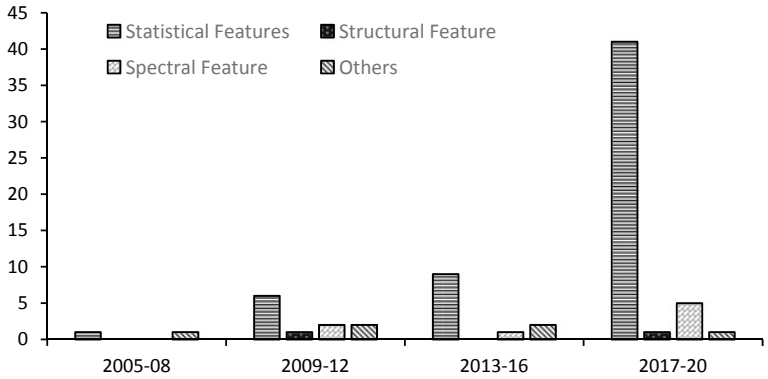


Fig. 2 Texture feature extraction techniques

Table 3 Feature extraction techniques based on color features

Techniques	Description
Histogram-based feature extraction techniques	A histogram contains the first order statistical information about the image. It can be used for the identification of plant diseases in two ways. Either its statistical properties like mean, standard deviation, variance, skewness, and kurtosis are used as features, or the entire histogram can be considered as a feature using intersection or discrimination [22, 33, 37, 39, 42, 55, 85, 94]
MPEG7 color feature-based techniques	MPEG7 color features are a collection of color descriptors that can be used to classify images based on colors. Several descriptors are used such as dominant color descriptor, scalable color descriptor, group of image or group of frames color descriptor, color layout descriptor, and color structure descriptor [97]
Color coherence vector	This technique computes whether a pixel is coherent or incoherent. This technique is based on the coherence of pixels that measure the degree to which the pixel is part a substantial large connecting region that has same color as the pixel being tested for coherence [33, 37, 39, 42, 55]
Color moments	In this technique, color moments are used as color features. Raw pixel intensity is used to calculate the moments like mean, standard deviation, skewness, kurtosis, and other higher moments [3, 6, 10, 13, 14, 17, 19, 21, 38, 41, 44, 50, 52, 54, 69, 78–80, 98–101]
Color co-occurrence matrix (CCM)	The input RGB images are converted to HSV color space, and then six color coherence matrixes are computed from each combination of color channels. Once these six matrixes are computed, second-order statistical attributes like variance, contrast correlation, etc., are calculated and used as features [5, 6, 24, 25, 27, 30–32, 35, 41, 46, 48, 51, 64, 84, 86–88, 102]
Other techniques	In such techniques, the pixel intensities are directly used to extract color features [4, 7, 34, 76, 80, 98, 103]

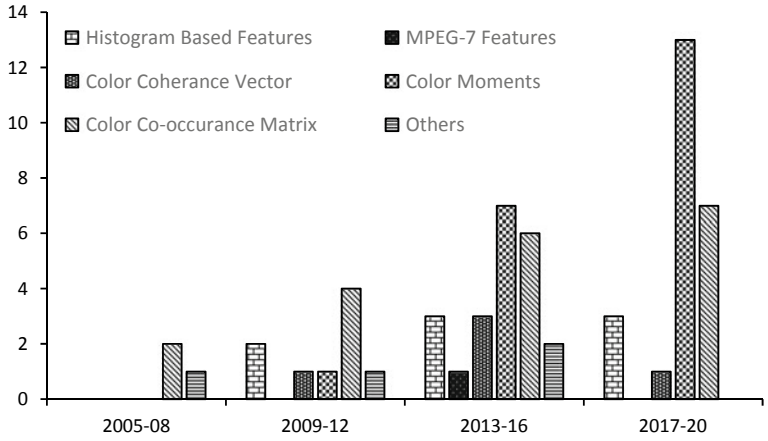


Fig. 3 Color feature extraction techniques

Table 4 Feature extraction techniques based on shape features

Techniques	Description
Geometric shape feature-based techniques	In these feature extraction techniques, geometric characteristics are used to extract shape features from an image. There are a variety of geometric shape features which include center of gravity or centroid, axis of minimal inertia, average binding energy, eccentricity, circularity ratio, ellipticity, rectangularity, solidity, and profiles [3, 7, 8, 10, 14, 15, 19, 21, 26, 38, 54, 72, 76, 79, 98, 104]
Moment invariant-based techniques	It is a very popular technique that uses moment invariants to describe a shape. Various moment invariants are used to extract shape-related information from an image such as contour moments, geometric moments, Zernike moment, radial Chebyshev moment, Legendre moment, homocentric polar-radius moment, orthogonal Fourier-Mellin moment, and pseudo Zernike moment [3, 13, 22, 42, 55, 85, 105]
Other techniques	Apart from the above, some other limited applications of feature extraction techniques have been reported based on polygon approximation shape features, scale space shape features, shape transform domain shape feature, and SIFT techniques [11, 39, 62, 78, 94, 106]

on color moments are mostly used as a popular and effective one. A gradual rise in the application of such techniques is also observed from 2005. Though the application of extraction techniques based on geometric shape features is gradually increasing, as a whole, shape features had the minimum application among the three groups of features: color, shape, and texture. Among all the feature extraction techniques, texture feature-based extraction techniques specifically using statistical approach

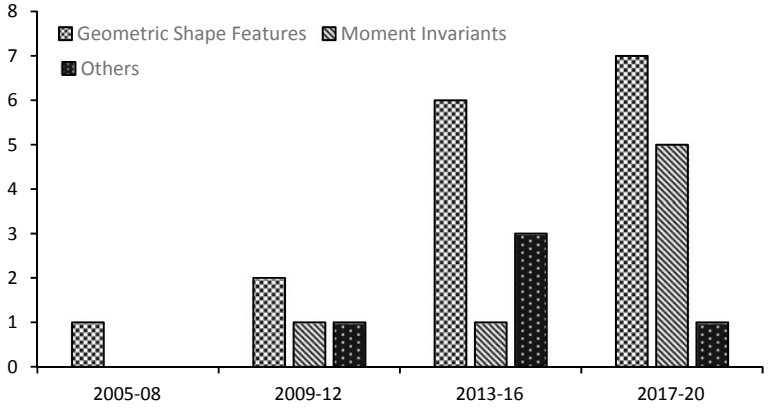


Fig. 4 Shape feature extraction techniques

have the highest application in designing plant disease diagnosis systems. A sudden rise in the application of such systems is observed from 2017 to 2020.

After the features are extracted, SVM has been applied to a maximum number of systems as a successful classifier to design image-based plant disease diagnosis systems. The application of SVMs has gradually increased since 2005. The application of ANN occupied the second position in designing such systems. During 2017–2020, the k-NN classifiers have some applications. The other classifiers have no remarkable applications.

5 Conclusion

The application of image processing techniques is gradually gaining momentum in several domains, including medical diagnosis, plant disease diagnosis, product grading, etc. This paper clarifies the past and present trend of applying such techniques in agricultural plant disease diagnosis since 2005. A total of 250 papers have been accessed that are published worldwide in reputed international journals on the related topics. Out of these 250 publications, 99 papers have been found contributory in the domain of image-based plant disease diagnosis in agriculture.

The application trends of the sub-techniques used in image processing such as segmentation, feature extraction, and machine learning have been studied in detail. This gives a clear view and covers all aspects of the application trend of such systems. The observation emphasizes that the application of image processing techniques in agricultural plant disease diagnosis will be an emerging research area in the near future. To date, no such comprehensive statistical study of application trend has been reported. This study will be very beneficial for ambitious researchers in the future to design such systems.

Table 5 Machine learning techniques

Techniques	Description
SVM	Support vector machine (SVM) is a popular supervised machine learning technique. In the learning phase, SVM learns the importance of each training data and estimates the decision boundary between output classes. The prediction of a data point in the test set is performed based on its distance to support vectors and the importance of support vectors. It can avoid over-fitting the data [2–5, 7, 9, 11, 12, 14, 17, 18, 20–22, 26, 33, 34, 36–39, 41–45, 48, 49, 51, 53–58, 61, 62, 64–70, 75, 76, 80, 82, 83, 88, 91–93, 98, 100–102, 105, 107]
ANN	Artificial neural network (ANN) is a very popular model used for classification and pattern recognition. The input of an ANN is a set of features extracted in the feature extraction phase and gives the output as identified disease(s). ANN is capable of learning the underlying relationships between the input features and the output classes [1, 13, 15, 19, 24, 29–32, 35, 40, 46, 49, 50, 52, 62, 64, 72, 74, 75, 79, 87, 90, 94, 97, 101, 102, 104, 108, 109]
k-NN	K nearest neighbor (k-NN) is one of the simplest supervised machine learning techniques. The training phase contains just storing the training dataset. In the testing phase, the prediction is performed by finding the nearest neighbor of the test data. This nearest neighbor is the training data point which is closest to the test data point [6, 8, 10, 16, 22, 24, 49, 52, 62, 63, 70, 81, 84, 89, 92, 95, 102]
Naïve Bayes classifier	Naïve Bayes classifier combines naïve Bayes probability model with a decision rule like selecting the highest probable hypothesis. In this classifier, the conditional probability of each feature is multiplied [110]. The class labeled returned for which the product obtained is largest [4, 22, 27, 28, 47, 49, 62, 73, 78]
Discriminant analysis	Discriminant analysis (DA) is capable of performing classification of an observation. It estimates the posterior probability of which the observation belongs to every class. Prior probabilities of class membership and estimation of multivariate normal distribution of each class with squared Mahalanobis distance is used to obtain the posterior probability [24, 25, 49, 86]
Decision tree	Decision tree is a multi-stage classifier. It learns the hierarchy of questions that are if/else in nature and then reach to the final decision [111]. In decision tree, each branch represents a test on an attribute, the intermediate nodes represent the output of the test, and the leaf nodes represents the output classes [22, 70]
Other classifiers	Apart from these classifiers, there are a few that are used like random forest, AdaBoost, logistic regression etc. [22, 49, 59, 70, 82, 85, 112]

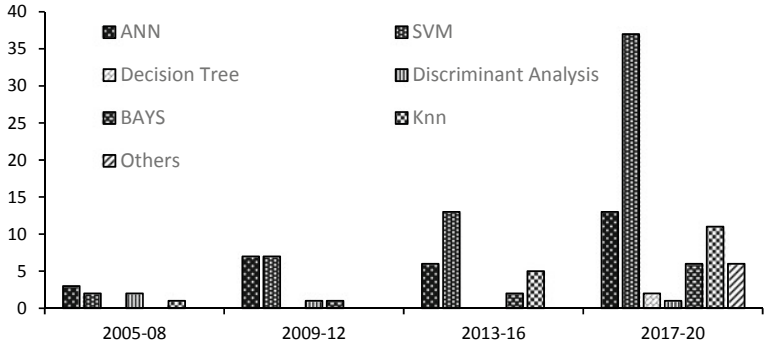


Fig. 5 Trend of machine learning techniques

References

1. Wang H, Li G, Ma Z, Li X (2012) Image recognition of plant diseases based on principal component analysis and neural networks. In: International conference on natural computation 2012, ICNC. IEEE, China, pp 246–251
2. Camargo A, Smith JS (2009) Image pattern classification for the identification of disease causing agents in plants. Comput Electron Agric 66(2):121–125
3. Tian Y, Zhao C, Lu S, Guo X (2011) Multiple classifier combination for recognition of wheat leaf diseases. Intell Autom Soft Comput 17(5):519–529
4. Phadikar S, Sil J, Das A (2012) Classification of rice leaf diseases based on morphological changes. Int J Inf Electron Eng 2(3):460–463
5. Arivazhagan S, Shebiah R, Ananthi S, Varthini S (2013) Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. Agric Eng Int CIGR J 15(1):211–217
6. Eaganathan U, Sophia J, Vinukumar L, Feroze J (2014) Identification of sugarcane leaf scorch diseases using K-means clustering segmentation and K-NN based classification. Int J Adv Comput Sci Technol (IJACST) 3(12):11–16
7. Patil S, Zambre R (2016) Classification of cotton leaf spot disease using support vector machine. Int J Eng Res 3(4):1511–1514
8. Shrivastava S, Hooda D (2014) Automatic brown spot and frog eye detection from the image captured in the field. Am J Intell Syst 4(4):131–134
9. Mokhtar U, El-Bendary N, Hassenian A, Emary E, Mahmoud M, Hefny H, Tolba M (2015) SVM-based detection of tomato leaves diseases. In: Advances in intelligent systems and computing. Springer, Switzerland, pp 641–652
10. Zhang SW, Shang YJ, Wang L (2015) Plant disease recognition based on plant leaf image. J Anim Plant Sci 25(3):42–45
11. Dandawate Y, Kokare R (2015) An automated approach for classification of plant diseases towards development of futuristic decision support system in Indian perspective. International conference on advances in computing, communications and informatics 2015, ICACCI. IEEE, India, pp 794–799
12. Islam M, Dinh A, Wahid K, Bhowmik P (2017) Detection of potato diseases using image segmentation and multiclass support vector machine. In: Canadian conference on electrical and computer engineering 2017, CCECE. IEEE, Canada, pp 1–4
13. Shao Y, Xuan G, Zhu Y, Zhang Y, Peng H, Liu Z, Hou J (2017) Research on automatic identification system of tobacco diseases. Imaging Sci J 65(4):252–259

14. Sharif M, Khan M, Iqbal Z, Azam M, Lali M, Javed M (2018) Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. Comput Electron Agric 150:220–234
15. Zhu J, Wu A, Wang X, Zhang H (2019) Identification of grape diseases using image analysis and BP neural networks. Multimedia Tools Appl 1–13
16. Mathur M (2019) Analysis of various plant disease detection techniques using KNN classifier. Int J Comput Sci Mobile Comput 8(7):65–70
17. Joshi P, Saha A (2019) Detection and classification of plant diseases using soft computing techniques. International conference on recent trends on electronics, information, communication & technology 2019, RTEICT. IEEE, India, pp 1289–1293
18. Poornima S, Kavitha S, Mohanavalli S, Sripriya N (2019) Detection and classification of diseases in plants using image processing and machine learning techniques. In: AIP Conf Proc 2095:030018
19. Rothe P, Rothe J (2019) Intelligent pattern recognition system with application to cotton leaf disease identification. In: Innovations in computer science and engineering. Springer Nature, Singapore, pp 19–27
20. Singh K, Kumar S, Kaur P (2019) Support vector machine classifier based detection of fungal rust disease in pea plant (*Pisam sativam*). Int J Inf Technol 11(3):485–492
21. Paul P, Aiswarya S, Aiswarya J, Athira S, Anagha C (2019) SasyaSneha—an approach for plant leaf disease detection. Int J Adv Trends Comput Sci Eng 8(5):2101–2106
22. Iqbal MdA, Talukder K (2020) Detection of potato disease using image segmentation and machine learning. In: International conference on wireless communications signal processing and networking 2020, WiSPNET. IEEE, India, pp 43–47
23. Ullah MdR, Hasnat A, Dola N (2019) Design and implementation of smart agriculture system and plant disease recognition. Daffodil International University
24. Pydipati R, Burks TF, Lee WS (2005) Statistical and neural network classifiers for citrus disease detection using machine vision. Trans ASAE 48(5):2007–2014
25. Pydipati R, Burks TF, Lee WS (2006) Identification of citrus disease using color texture features and discriminant analysis. Comput Electron Agric 52(1–2):49–59
26. Yao Q, Guan Z, Zhou Y, Tang J, Hu Y, Yang B (2009) Application of support vector machine for detecting rice diseases using shape and color texture features. In: International conference on engineering computation 2009, ICEC. IEEE, China, pp 79–83
27. Bandi S, Varadharajan A, Chinnasamy A (2013) Performance evaluation of various statistical classifiers in detecting the diseased citrus leaves. Int J Eng Sci Technol 5(2):298–307
28. Mondal D, Kole D, Roy K (2017) Gradation of yellow mosaic virus disease of okra and bitter gourd based on entropy based binning and Naive Bayes classifier after identification of leaves. Comput Electron Agric 142:485–493
29. Dhakate M, Ingole AB (2015) Diagnosis of pomegranate plant diseases using neural network. In: National conference on computer vision, pattern recognition, image processing and graphics 2015, NCVPRIPG. IEEE, India, pp 1–4
30. Al Bashish D, Braik M, Bani-Ahmad S (2010) A framework for detection and classification of plant leaf and stem diseases. International conference on signal and image processing 2010, ICSIP. IEEE, India, pp 113–118
31. Al-Hiary H, Bani-Ahmad S, Reyalat M, Braik M, Alrahamneh Z (2011) Fast and accurate detection and classification of plant diseases. Int J Comput Appl 17(1):31–38
32. Al Bashish D, Braik M, Bani-Ahmad S (2014) Detection and classification of leaf diseases using K-means-based segmentation and neural-networks-based classifier. Inf Technol J 10(2):4638–4645
33. Dubey S, Jalal A (2012) Detection and classification of apple fruit diseases using complete local binary patterns. International conference on computer and communication technology 2012, ICCCT. IEEE, India, pp 346–351
34. Sannakki S, Rajpurohit V, Nargund VB (2013) SVM-DSD: SVM based diagnostic system for the detection of pomegranate leaf diseases. In: Proceedings of international conference on advances in computing 2013. Springer, India, pp 715–720

35. Sannakki S, Rajpurohit V, Nargund VB, Kulkarni P (2013) Diagnosis and classification of grape leaf diseases using neural networks. In: International conference on computing, communications and networking technologies 2013, ICCCNT. IEEE, India, pp 1–5
36. Gavhale K, Gawande U, Hajari K (2014) Unhealthy region of citrus leaf detection using image processing techniques. In: International conference for convergence for technology 2014, ICCT. IEEE, India, pp 1–6
37. Dubey S, Jalal A (2012) Adapted approach for fruit disease identification using images. Int J Comput Vis Image Process 2(3):44–58
38. Pujari J, Yakkundimath R, Byadgi A (2014) Detection and classification of fungal disease with radon transform and support vector machine affected on cereals. Int J Comput Vis Robot 4(4):261–280
39. Bhange M, Hingoliwala HA (2015) Smart farming: pomegranate disease detection using image processing. Proc Comput Sci 58:280–288
40. Awate A, Deshmankar D, Amrutkar G, Bagul U, Sonavane S (2015) Fruit disease detection using color, texture analysis and ANN. International conference on green computing and internet of things 2015, ICGCIoT. IEEE, India, pp 970–975
41. Padol P, Yadav A (2016) SVM classifier based grape leaf disease detection. Conference on advances in signal processing 2016, CASP. IEEE, India, pp 1–6
42. Dubey S, Jalal A (2016) Apple disease classification using color, texture and shape features from images. SIViP 10(5):819–826
43. Masazhar A, Kamal M (2017) Digital image processing technique for palm oil leaf disease detection using multiclass SVM classifier. International conference on smart instrumentation, measurement and application 2017, ICSIMA. IEEE, Malaysia, pp 1–6
44. Agrawal N, Singhai J, Agarwal D (2017) Grape leaf disease detection and classification using multi-class support vector machine. International conference on recent innovations in signal processing and embedded systems 2017, RISE. IEEE, India, pp 138–244
45. Habib Md, Majumder A, Jakaria A, Akter M, Uddin M, Ahmed F (2018) Machine vision based papaya disease recognition. J King Saud Univ Comput Inf Sci
46. Sreedevi S, Bhashith T, Pushpa R (2019) Detection and classification of plant leaf diseases using image processing. Int J Latest Trans Eng Sci (IJLTES) 7(2):028–031
47. Prema P, Veeramani A, Theradimani M, Sivakumar T (2019) Plant leaf disease detection using Curvelet transform. J Pharmacognosy Phytochem 8(5):2314–2317
48. Banni R, Gorabal J (2019) A novel based approach for detection of canker disease on citrus leaves using image processing methods. In: International conference on intelligent computing and communication technologies. Springer Nature, Singapore, pp 322–327
49. Verma T, Dubey S (2019) Impact of color spaces and feature sets in automated plant diseases classifier: a comprehensive review based on rice plant images. Arch Comput Methods Eng 1–22
50. Chanda M, Biswas M (2019) Plant disease identification and classification using back-propagation neural network with particle swarm optimization. In: 3rd international conference on trends in electronics and informatics 2019, ICOEI. IEEE, pp 1029–1036
51. Jenifa A, Ramalakshmi R, Ramachandran V (2019) Classification of cotton leaf disease using multi-support vector machine. In: International conference on intelligent techniques in control, optimization and signal processing 2019, INCOS. IEEE, pp 1–4
52. Ramesh S, Vydeki D (2019) Application of machine learning in detection of blast disease in south Indian rice crops. J Phytology 11(1):31–37
53. Agarwal A, Sarkar A, Dubey A (2019) Computer vision-based fruit disease detection and classification. In: Smart innovations in communication and computational sciences. Springer Nature, Singapore, pp 105–115
54. Alehegn E (2019) Ethiopian maize diseases recognition and classification using support vector machine. Int J Comput Vis Robot 9(1):90–109
55. Ayyub S, Manjramkar A (2019) Fruit disease classification and identification using image processing. In: 3rd international conference on computing methodologies and communication 2019, ICCMC. IEEE, pp 754–758

56. Wahab A, Zahari R, Lim T (2019) Detecting diseases in chili plants using K-means segmented support vector machine. In: 3rd international conference on imaging, signal processing and communication 2019, ICISPC. IEEE, pp 57–61
57. Chawal B, Panday S (2019) Rice plant disease detection using twin support vector machine (TSVM). J Sci Eng 7:61–69
58. Kaur N, Kaur P (2019) Classification and segmentation approach for plant disease detection. Int J Comput Sci Mob Comput 8(8):6–16
59. Devaraj A, Rathan K, Jaahnavi S, Indira K (2019) Identification of plant disease using image processing technique. In: International conference on communication and signal processing 2019, ICCSP. IEEE, pp 0749–0753
60. Dharmasiri S, Jayalal S (2019) Passion fruit disease detection using image processing. In: International research conference on smart computing and systems engineering 2019, SCSE. IEEE, pp 126–133
61. Sethy P, Barpanda N, Rath A (2019) Detection and identification of rice leaf diseases using multiclass SVM and particle swarm optimization technique. Int J Innov Technol Explor Eng (IJITEE) 8(6S2):108–120
62. Devi G, Neelamegam P (2019) Image processing based rice plant leaves diseases in Thanjavur, Tamil Nadu. Cluster Comput 22(6):13415–13428
63. Hossain E, Hossain F, Rahaman M (2019) A color and texture based approach for the detection and classification of plant leaf disease using KNN classifier. In: International conference on electrical, computer and communication engineering 2019, ECCE. IEEE, pp 1–6
64. Vamsidhar E, Rani J, Babu R (2019) Plant disease identification and classification using image processing. Int J Eng Adv Technol (IJEAT) 8(3S):442–446
65. Latha M, Jaya S (2019) Detection of Septoria spot on blueberry leaf images using SVM classifier. ICTACT J Image Video Process 9(4):2015–2019
66. Santhosh A, Saravanakumar K, Saravanakumar P, Vimalraj S, Saranya R (2019) Detection of leaf diseases and classifying them using multiclass SVM. Int Res J Eng Technol (IRJET) 6(3):421–426
67. Butale N, Kodavade D (2019) Detection of plant leaf diseases using image processing and soft-computing techniques. Int Res J Eng Technol (IRJET) 6(6):3288–3291
68. Kuricheti G, Supriya P (2019) Computer vision based turmeric leaf disease detection and classification: a step to smart agriculture. In: 3rd international conference on trends in electronics and informatics 2019, ICOEI. IEEE, pp 545–549
69. DSpace Repository. http://dspace.daffodilvarsity.edu.bd:8080/handle/123456789/3700. Last accessed 10 Dec 2019
70. Panchal P, Raman V, Mantri S (2019) Plant diseases detection and classification using machine learning models. In: 4th international conference on computational systems and information technology for sustainable solution 2019, CSITSS, vol 4. IEEE, pp 1–6
71. Jaypee University of Information and Technology. http://www.ir.juit.ac.in:8080/jspui/handle/123456789/22879. Last accessed 10 Dec 2019
72. Bansod V (2019) Rice crop disease identification and classifier. Int J Comput Sci Eng 7(11):45–48
73. Kaur D (2019) Textual feature analysis and classification method for the plant disease detection. Int J Comput Sci Mob Comput 8(9):147–154
74. Huang K (2007) Application of artificial neural network for detecting Phalaenopsis seedling diseases using color and texture features. Comput Electron Agric 57(1):3–11
75. Meunkaewjinda A, Kumsawat P, Attakitmongcol K, Srikaew A (2008) Grape leaf disease detection from color imagery using hybrid intelligent system. In: International conference on electrical engineering/electronics, computer, telecommunications and information technology 2008, ECTI-CON, vol 1. IEEE, Thailand, pp 513–516
76. Youwen T, Tianlai L, Yan N (2008) The recognition of cucumber disease based on image processing and support vector machine. In: Congress on image and signal processing 2008, CISP, vol 2. IEEE, China, pp 262–267

77. Guru DS, Mallikarjuna PB, Manjunath S (2011) Segmentation and classification of tobacco seedling diseases. In: Proceedings of the fourth annual ACM Bangalore conference 2011, pp 1–5
78. Phadikar S, Sil J, Das A (2013) Rice diseases classification using feature selection and rule generation techniques. Comput Electron Agric 90:76–85
79. Aji A, Munajat Q, Pratama A, Kalamullah H, Setiyawan J, Arymurthy A (2013) Detection of palm oil leaf disease with image processing and neural network classification on mobile device. Int J Comput Theory Eng 5(3):528
80. Han L, Haleem M, Taylor M (2015) A novel computer vision-based approach to automatic detection and severity assessment of crop diseases. In: Science and information conference 2015, SAI. IEEE, UK, pp 638–644
81. Parikh A, Raval M, Parmar C, Chaudhary S (2016) Disease detection and severity estimation in cotton plant from unconstrained images. In: IEEE international conference on data science and advanced analytics 2016, DSAA. IEEE, Canada, pp 594–601
82. Jaisakthi S, Mirunalini P, Thenmozhi D, Vatsala (2019) Grape leaf disease identification using machine learning techniques. In: International conference on computational intelligence in data science 2019, ICCIDS. IEEE, pp 1–6
83. Veling S, Kalelkar R, Ajgaonkar L, Mestry N, Gawade N (2019) Mango disease detection by using image processing. Int J Res Appl Sci Eng Technol (IJRASET) 7(4):3717–3726
84. Vaishnnave MP, Devi S, Srinivasan P, Arut Perum Jothi G (2019) Detection and classification of groundnut leaf diseases using KNN classifier. In: International conference on system, computation, automation and networking 2019, ICSCAN. IEEE, pp 1–5
85. Govardhan M, Veena MB (2019) Diagnosis of tomato plant diseases using random forest. In: Global conference for advancement in technology, GCAT. IEEE, pp 1–5
86. Kim D, Burks T, Qin J, Bulanon D (2009) Classification of grapefruit peel diseases using color texture feature analysis. Int J Agric Biol Eng 2(3):41–50
87. Ramakrishnan M (2015) Groundnut leaf disease detection and classification by using back propagation algorithm. International conference on communications and signal processing 2015, ICCSP. IEEE, India, pp 0964–0968
88. Sood A, Singh P (2020) Hybrid system for detection and classification of plant disease using qualitative texture features analysis. Proc Comput Sci 167:1056–1065
89. Prasad S, Peddoju S, Ghosh D (2016) Multi-resolution mobile vision system for plant leaf disease diagnosis. SIViP 10(2):379–388
90. Kai S, Zhikun L, Hang S, Chunhong G (2011) A research of maize disease image recognition of corn based on BP networks. In: International conference on measuring technology and mechatronics automation 2011, vol 1. IEEE, China, pp 246–249
91. Asraf H, Nooritawati MT, Rizam MSB (2012) A comparative study in kernel-based support vector machine of oil palm leaves nutrient disease. Proc Eng 41:1353–1359
92. Deshapande A, Giraddi S, Karibasappa KG, Desai S (2019) Fungal disease detection in maize leaves using Haar wavelet features. In: Information and communication technology for intelligent systems. Springer, pp 275–286
93. Khan M, Uandai S, Srinivasan H (2019) Anthracnose disease diagnosis by image processing, support vector machine and correlation with pigments. J Plant Pathol 101(3):749–751
94. Xu G, Zhang F, Shah S, Ye Y, Mao H (2011) Use of leaf color images to identify nitrogen and potassium deficient tomatoes. Pattern Recogn Lett 32(11):1585–1590
95. Aruraj A, Alex A, Subathra MSP, Sairamya NJ, George T, Ewards V (2019) Detection and classification of diseases of banana plant using local binary pattern and support vector machine. In: 2nd international conference on signal processing and communication 2019, ICSPC. IEEE, pp 231–235
96. Chaki J, Dey N (2020) Texture feature extraction techniques for image recognition. Springer Nature, Singapore
97. Molina J, Gil R, Bojaca C, Gomez F, Franco H (2014) Automatic detection of early blight infection on tomato crops using a color based classification strategy. In: Symposium on image, signal processing and artificial vision 2014. IEEE, Colombia, pp 1–5

98. Chung C, Huang K, Chen S, Lai M, Chen Y, Kuo Y (2016) Detecting Bakanae disease in rice seedlings by machine vision. Comput Electron Agric 121:404–411
99. Sabrol H, Kumar S (2016) Intensity based feature extraction for tomato plant disease recognition by classification using decision tree. Int J Comput Sci Inf Secur 14(9):622–626
100. Hossain M, Mou R, Hasan M, Chakraborty S, Razzak M (2018) Recognition and detection of tea leaf's diseases using support vector machine. International colloquium on signal processing & its applications 2018, CSPA. IEEE, Malaysia, pp 150–154
101. Kumari U, Vignesh A, Panigrahy A, Ramya L, Padma T (2019) Int J Innov Technol Explor Eng 8(10):3364–3367
102. Chang Y, Mahmud Md, Shin J, Nguyen-Quang T, Price G, Prithiviraj B (2019) Comparison of image texture based supervised learning classifiers for strawberry powdery mildew detection. AgriEngineering 1(3):90–109
103. Chaki J, Dey N (2020) Image color feature extraction techniques: fundamentals and applications. Springer Nature, Singapore
104. Zhu J, Wu A, Wang X, Zhang H (2020) Identification of grape diseases using image analysis and BP neural networks. Multimedia Tools Appl 79(21):14539–14551
105. Kaur P, Pannu H, Malhi A (2019) Plant disease recognition using fractional-order Zernike moments and SVM classifier. Neural Comput Appl 31(12):8749–8768
106. Chaki J, Dey N (2019) A beginner's guide to image shape feature extraction techniques. CRC Press
107. Jian Z, Wei Z (2010) Support vector machine for recognition of cucumber leaf diseases. In: International conference on advanced computer control 2010, vol 5. IEEE, China, pp 164–266
108. Basheer I, Hajmeer M (2000) Artificial neural networks: fundamentals, computing, design, and application. J Microbiol Methods 43(1):3–31
109. Guru DS, Mallikarjuna PB, Manjunath S (2011) Segmentation and classification of tobacco seedling diseases. In: Proceedings of the fourth annual ACM Bangalore conference 2011. ACM, India, pp 1–5
110. Stegmayer G, Milone D, Garran S, Burdyn L (2013) Automatic recognition of quarantine citrus diseases. Expert Syst Appl 40(9):3512–3517
111. Safavian S, Landgrebe D (1991) A survey of decision tree classifier methodology. IEEE Trans Syst Man Cybern 21(3):660–674
112. Muller A, Guido S (2016) Introduction to machine learning with Python: a guide for data scientists. O'Reilly Media Inc., CA

A Bibliometric Analysis of E-Governance and ICT

Smita Wagholikar and Arti Chandani

Abstract The objective of this paper is to study the trends in the research papers in last 2 decades in the area of e-governance and ICT. The present study is aimed to find the existing literature in the domain of e-governance and ICT by undertaking a bibliometric review. This work examines the research papers which were extracted from Scopus database. Scopus is one of the largest databases and it is much user friendly in terms of extracting the existing studies undertaken by other researchers. The papers extracted from Scopus were from year 2001 to 2022 and authors have used only one filter in addition to year, i.e. language as English. The authors have used VOSviewer software to generate network diagrams. The results show that there has been a significant increase in the number of research studies since 2005 onwards. Around 47.3% of research studies are published in the conference proceedings and there are around 28.5% as journal articles. Another important finding was that India was at the top of the list in terms of number of article by a country where India has 229 research papers in the area of e-governance and ICT. The study also gives an insight into the co-authorship and keyword analysis. The study is very important from the perspective that it gives a direction to the future researchers by apprising the existing studies in the area of e-governance and ICT. The study is unique in a way that it gives co-authorship and keyword analysis using VOSviewer.

Keywords E-governance · ICT · VOSviewer · Scopus · Network diagram

1 Introduction

Information and communication technology (ICT) has established its importance and application in various sectors, including that of governance. The utilization of ICT to

S. Wagholikar (✉)
SIT, Symbiosis International University, Pune, Maharashtra 411028, India
e-mail: smita.wagholikar@sitpune.edu.in

A. Chandani
SIMS, Symbiosis International University, Pune, Maharashtra 411028, India

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_67

expedite services and communication to the government is referred to as electronic governance or simply e-governance. Employing the ICT in government services to all citizens at local, regional and national level helps attain transparency and faster communication. The e-governance services can be designed keeping the requirements of the public, simplifying the processes to save time, increases accessibility and thereby encouraging the efficient use of ICT in e-governance. This increases the involvement of citizens in government processes and helps disseminate information effectively.

There is inevitably a need to ensure that there is adequate awareness and use of the facilities provided by the government to fasten the transactions, revenue collection, taxes paid by the public and communication or grievance handling through e-governance. Also, a more focused effort is required to bridge the gap between the higher and lower classes of the society and e-governance can do so. To promote this objective, Rajeshwar and Roy [12] reported the successful implementation of "Lokvani". This service not only reached to all the sectors and strata of the society by providing solution to problems but also was found to lessen the number of visitors to government departments.

2 Literature Review

The existence of ICT in various sections of governance in our country is over two decades by Srivastava and Sharma [15]. However, the awareness and acceptance of ICT in e-governance over the last decade has increased substantially as seen from the literature (Table 1). The facility of e-governance needs to be made accessible to all irrespective of the financial status and urban or rural area of residence. Besides, the empathetic approach to complaints and maintaining transparency in processes in order to provide quality service to all class of people highlights the benefits of ICT in e-governance. The policy of introducing the Unique Identity of India UID-Aadhaar

Table 1 Publication count between years 2001–2022

Year	Count	Year	Count	Year	Count
2001	1	2008	39	2015	45
2002	2	2009	36	2016	45
2003	6	2010	39	2017	42
2004	4	2011	56	2018	33
2005	16	2012	46	2019	50
2006	20	2013	28	2020	38
2007	38	2014	55	2021	20
				2022	1
Total					660

was taken by the Indian government. With an intention of issuing a genuine and digital identity for rendering easy access to many government services, the Aadhaar was created and readily accepted by one and all. This initiative helped gain transparency, social justice and avoid any malpractices [15]. At an initial stage implementation of ICT in e-governance, there were serious issues related to ethics, security and cooperation of professionals [9]. These issues were handled by policy makers.

A survey by Boksova et al. [3] mentions that there exists a gap between Internet users and general public regarding the digitization of public services. Although the situation is not a serious challenge, consistent improvement is carried out to meet the demands, satisfy citizens and encourage the use of digital services for e-governance. The research identifies that citizens above the age of sixty and citizens with lesser economic status showed insecurity and were less involved towards the digitization process. Keeping this observation in mind, the government established help centres at various locations accessible to people. Gupta et al. [5] put forth the concept of setting–up of tele-centre to make information and communication technology infrastructure accessible to people in rural areas and benefit all residents for overall development. However, there are challenges such as deciding on location of centres and their sustainability. Khan et al. [7] discusses the identification of e-satisfaction parametres of citizens to ensure successful e-governance in Pakistan by setting up of a "Union Committee Help Desk".

A study by Bhuvana and Vasantha [2] gauges the success of e-governance in India by conducting a survey of rural section of Tiruppur district. Results show that citizens are convinced about the use of e-services but face problems in accessing the ICT services. It is evident that there is a scope to understand the requirements and take necessary measures to involve a large number of people to use e-governance services. Patnaik et al. [11] report the use of latest technology like artificial intelligence and machine learning in e-governance to make it advanced and add to the national development. The researchers study the various records in Odisha and report the use of big data, data analytics for better results and sustainability. Babu et al. [1] report the issue of digital divide that needs to be cross-cut to reduce socio-economic differences. The facility of mobile phone has come to the rescue and ICT is now more accessible to the deprived and less literate at an affordable price. Alirajpur district (Madhya Pradesh in India) was considered for the study. The positive results indicate that ICT is becoming more reachable for achieving e-governance. Mili [10] reported the government initiative of applying ICT in the development of rural economy in Assam, India with reasonably good success.

An overview by Kumar et al. [8] described the recent updates and progress in the use of ICT for holistic development and socio-economic uplift in the rural India. The research mentions the housing scheme for rural people through digitization. This scheme is under application on a wider scale so as to benefit a larger number of beneficiaries. Besides, initiatives and efforts for changing the rural scenario in India, the smart city initiatives towards generating a completely ICT-based infrastructure are reported by Sharma and Kalra [13]. The ICT supported e-governance provides a holistic makeover of cities to smart cities with availability of smart services. The

researchers also proposed a scheme the remote user scheme for e-governance in smart cities.

Jacob et al. [6] studied the study on two municipal corporations in rural areas of South Africa, wherein ICT-based e-governance was successful in increasing the capacity of work, management, access to information and led to effective maintenance of records. The awareness and implementation of ICT-based projects in Serbia led to a successful initiative for e-governance as reported by Damnjanović [4]. The study is said give a clear insight into the advantages and failures for adoption of ICT in Eastern Europe and improvement in executing the same. This clearly implies that the implementation of ICT in e-governance is proving its worth whether in rural or urban regions in India or in any other country. Sharma and Ray [14] suggested the use of mobile-based ICT in addition to computer-based model for executing initiatives for e-governance. This is likely to reduce the burden telecentres face and in turn increase the sustainability of the processes. There is also a need for the government to update their skills and technology for telecentres to meet the public demand.

2.1 Research Gap

The review of literature widely addressed the limited involvement of rural population in ICT-based e-governance. There is a need of systematic and careful analysis and review of e-governance and monitoring whilst encouraging the explicit use of ICT for governance. Also, to the best of our knowledge the literature does not reveal any bibliometric analysis for any study and reference.

2.2 Research Objectives

(A) The present research paper aims at analysing the use of ICT for facilitating e-governance for holistic development of rural and urban regions along with bibliometric analysis.
(B) To find co-authorship of authors and countries using network diagram.
(C) To find out author keywords, index keywords in the area of e-governance and ICT.

3 Research Design

The present research is aimed to present the current scenario of use of ICT in e-governance using bibliometric analysis. We have used the Scopus database for extracting the articles indexed in Scopus. The search word used was e-governance and ICT, in article title, abstract, keyword, which resulted in the articles 666. We

applied filters to extract the articles and one of the filters was the time period. The articles for the period of 2001 to 2022 were included in the research. The language "English" used another filter.

Using these filters, total articles were 666 which were downloaded from Scopus in csv format. Once the file was downloaded, we checked the articles and there were some missing entries and duplicate entries which were cleaned. So the final number of articles were 660. VOSviewer was the tool which was used by us in order to generate network diagrams and the csv file which was taken from Scopus was used as the source data for this. The articles were extracted on 21 September, 2021. The results of the network diagram are presented in Sect. 4 under the data analysis section.

4 Data Analysis

The present section deals with the analysis of the articles which were extracted from Scopus.

We have used VOSviewer to draw the network diagrams for the purpose of bibliometric analysis. Figure 1 shows the co-authorship of authors. There were 1154 authors of the data extracted from Scopus using the filters as given in the research design section. We have used 2 documents as the minimum number of documents and 1 citation as the minimum citation of an author. The network diagram shows that Malhotra C. has the strongest link amongst the selection with 11 as the total link strength and 4 documents which was followed by many authors who had total link strength of 10.

Figure 2 shows a network diagram of authors where the unit of analysis is country. The sample data has authors from 99 countries and we have used minimum 1 document and minimum 1 citation and the resultant diagram is shown in Fig. 2. The maximum link strength is India and USA with a total link strength of 24 each followed by 21 of the UK. If we see the maximum number of documents with respect to country, India has 229 documents followed by USA with 53 documents and Bangladesh with 36 documents whilst in terms of citations maximum citation are 636 for India, followed by 431 of USA and 391 of UK. This shows that in terms of e-governance and ICT the majority of the research is being carried out in India, USA and UK.

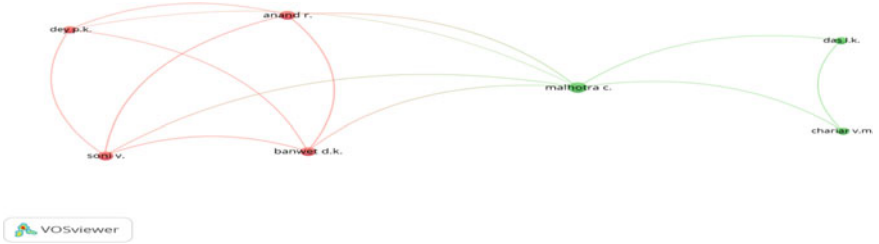


Fig. 1 Co-authorship of authors

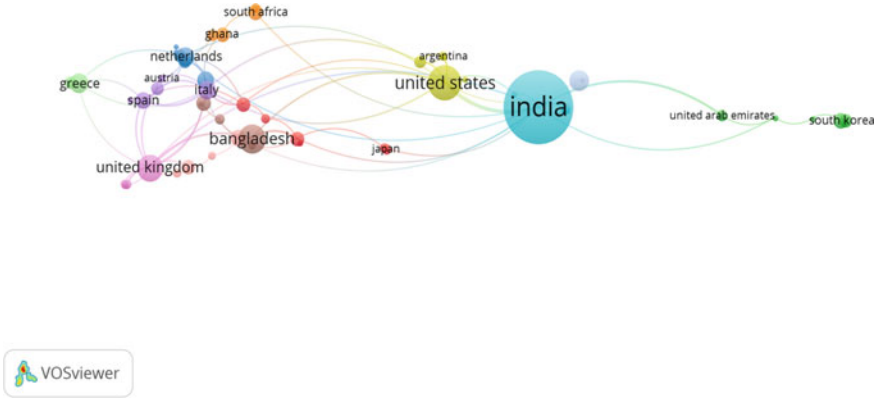


Fig. 2 Co-authorship of authors using country

The total number of keywords were 3042 for the sample data whilst we have used the 2 as the minimum number of occurrence of keywords which results in 702 keywords. The highest link strength was of the word "E-Governance" with 2839 followed by "Government Data Processing" as 2156. In terms of occurrence also "E-Governance" and "Government Data Processing" were high with a number of 377 and 231 occurrences, respectively. This also highlights that when researchers are working in the area of e-governance, the word e-governance and Government data processing is used much. The other keywords which are also being used by researchers are e-government and ICT amongst others (Fig. 3).

The further analysis of keywords was done as shown in Fig. 4 where we used the filter of minimum 2 author keywords which resulted in 295 keywords whilst the total keywords were 1464. The words like "E-Governance", "ICT" and "E-Government" had the maximum total link strength as 754, 345, 289, respectively. The occurrences of these author keywords were 285, 119 and 94, respectively, for "E-Governance",

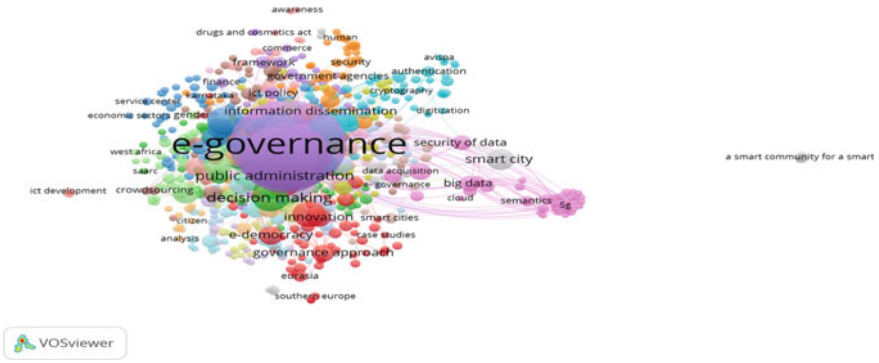


Fig. 3 Co-occurrence of all keywords

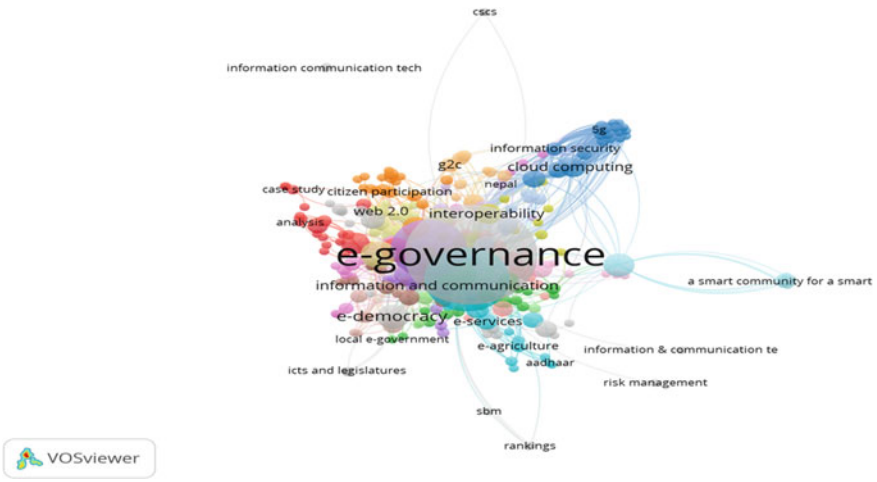


Fig. 4 Co-occurrence of author keywords

"ICT" and "E-Government". This shows the importance of the words in this given topic.

The index keywords are shown in the network diagram as shown in Fig. 5 where we have used the inclusion criteria as minimum of 2 index keywords. There were in total 2218 keywords but 535 met the threshold. "E-Governance", "Government

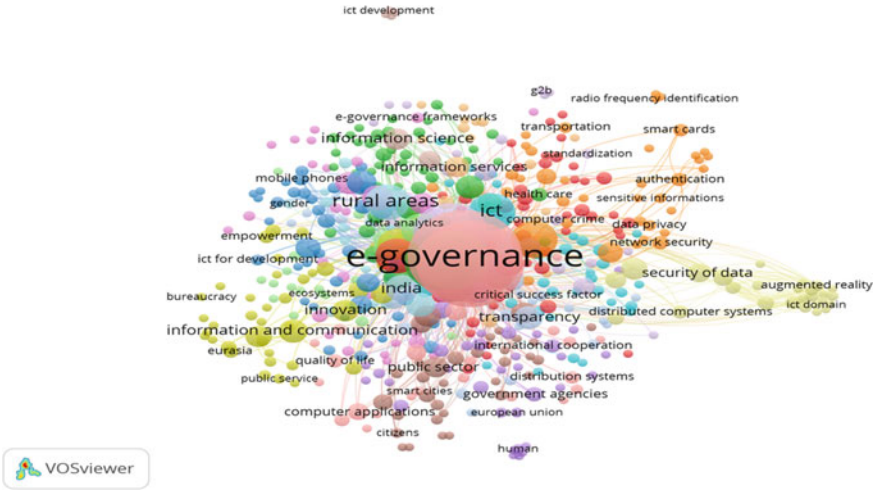


Fig. 5 Co-occurrence of index keywords

Data Processing", "Information and Communication Technologies" and "Information Technology" had the total link strength of 2161, 1937, 977, and 957, respectively, whilst the maximum occurrence was of "E-Governance" with a number of 260 followed by "E-Governance" at 231.

5 Conclusion

ICT has proved to be instrumental in changing the face of governance by approaching citizens in all economic strata as well as rural and urban regions. The digital literacy and technology can be made accessible by establishing help centres for citizen empowerment. The concept of smart cities and smart living can be achieved by using ICT in e-governance. Besides this, ICT in e-governance can be applied in agriculture, transport services and capacity building when executed by municipal corporations or public management systems. It helps in achieving transparent and speedy services in public administration.

5.1 Limitations

The study herein uses the database Scopus only for bibliometric analysis. There are other databases which, if used, will yield different results however Scopus is one of the largest database therefore we have used that. Secondly another limitation of the study emanates from the fact that we have used year, i.e. 2001–2022 as the time period and "English" as the language as the filters to extract the articles. The results and implications would be different if less or more filters are applied to the dataset. Thirdly, the limitation of the selected keywords become the limitation of the study. We have used "E-Governance" and "ICT" as the keywords to select the articles, however this was the objective of the study therefore this becomes the scope of the study. Lastly the software used to generate the network diagram was VOSviewer and there is other software available which might give a different result. Again we used VOSviewer as the software as it is one of the most widely used software for the bibliometric studies in the research.

5.2 Implications

The advent of ICT in e-governance and public administration opens the doors to uplift and progress of rural and urban domains and paves the way to smart living. Application of ICT in public management reduces stress on government offices and facilitates easy transparent services to citizens at a click on their computers or mobile phones. The aim of a healthy set up for technology driven decentralized effective

e-services for inclusive governance is achieved by using ICT in e-governance. The study is very useful from the perspective of research as well as from a managerial perspective. This study gives a landscape of the research in the area of e-governance and ICT along with highlighting the areas in which the research has not been done which can be taken up by the future researchers working in the field of e-governance and ICT.

5.3 Scope for Future Research

Although ICT has entered into various sections like agriculture, participative management and e-governance making services smooth; the application of ICT for sustainable development needs to be explored. The advantages and implementation of e-governance needs to be introduced to students in their curriculum. This will motivate students to design more convenient software for e-governance or start-ups and entrepreneurship. There can be more policies implemented for using e-governance which can be applied in various fields with the help of ICT.

References

1. Babu KR, Sagar AB, Kothari P (2020) Social inclusion and e-governance: a case study in Alirajpur district. In: International conference on innovations for community services. Springer, Cham, pp 285–294
2. Bhuvana M, Vasantha S (2020) Neural network machine learning techniques using R studio for predicting the attitude of rural people for accessing E-governance services. J Adv Res Dyn Control Syst 12(2):2552–2562
3. Boksova J, Boksa M, Horak J, Pavlica K, Strouhal J, Saroch S (2021) E-Government services and the digital divide: a quantitative analysis of the digital divide between the general public and internet users. J Telecommun Digital Econ 9(1):27–49
4. Damnjanović I (2019) Democratic innovations in Serbia: a misplaced trust in technology. Contemp Polit 25(1):111–127
5. Gupta R, Muttoo SK, Pal SK (2021) A comprehensive review of approaches used for solving tele-centre location allocation problem in geographical plane. Int J Oper Res 40(2):200–238
6. Jacobs C, Rivett U, Chemisto M (2019) Developing capacity through co-design: the case of two municipalities in rural South Africa. Inf Technol Dev 25(2):204–226
7. Khan K, Zaki H, Faizan A, Ali SA, Alvi U (2020) Elucidation of e-governance in Pakistan: a roadway to a Finer state. In: 2020 international conference on information science and communication technology (ICISCT). IEEE, pp 1–6
8. Kumar P, Mittal P, More A, Ahmed B (2020) ICT enabled implementation of rural housing scheme for sustainable rural development. In: ICT systems and sustainability. Springer, Singapore, pp 287–297
9. Lodge J (2007) Freedom, security and justice: the thin end of the wedge for biometrics? ANNALI-ISTITUTO SUPERIORE DI SANITA 43(1):20
10. Mili M (2019) Information technology and e-governance: comparative policy perspective of India and Assam. Int J Adv Sci Technol 28(20):819–822

11. Patnaik P, Pattnaik S, Singh P (2020) Use of data analytics for effective e-governance: a case study of "EMutation" system of Odisha. In: Advances in data science and management. Springer, Singapore, pp 109–122
12. Rajeshwar K, Roy SM (2022) Escalating social justice through effective implementation of e-governance. In: Information and communication technology for competitive strategies (ICTCS 2020). Springer, Singapore, pp 183–191
13. Sharma G, Kalra S (2019) Advanced multi-factor user authentication scheme for E-governance applications in smart cities. Int J Comput Appl 41(4):312–327
14. Sharma R, Ray S (2019) Exploring the impact of mobile technology adoption on financial sustainability of telecentres: the case of India. In: Proceedings of the 12th international conference on theory and practice of electronic governance, pp 427–434
15. Srivastava AK, Sharma S (2017) Social justice through Aadhaar: an e-policy initiative. In: Technology, society and sustainability. Springer, Cham, pp 83–97

Face Recognition for Surgical and Non-surgical Faces Using Patch-Based Approach

Kavita R. Singh and Roshni S. Khedgaonkar

Abstract Face Recognition (FR) is the task of remembering the individuals from their faces rather than any other body parts. In addition to all existing non-surgical challenging problems, Facial Plastic Surgery (FPS) is one of the most recent challenging issue to FR problems. Specifically, FPS is the medical process of restoration, reconstruction, or enhancing the appearance of some portion of the face. FPS causes significant difficulties since it permanently alters some face features, which alters the facial appearance further. These challenging issues of variations in facial appearances either due to non-surgical or surgical conditions motivate us to understand the effect of each on FR, respectively. The study of face recognition has been conducted in this chapter in non-surgical and surgical situations using a patch-based approach in which facial statistical features are extracted for recognition using a nearest neighbor concept such as a k-NN classifier. The resented study is tested on different facial datasets.

Keywords Face recognition · Facial plastic surgery · PSD · Facial dataset

1 Introduction

Face recognition techniques have proven themselves as a non-intrusive mechanism for personal identification. Due to this, face recognition techniques; have found its diverged usability in the surveillance system, crime controlling bodies, law enforcement applications, commercial purposes, child care centers, social networking sites, etc., However, FR still facing performance issues because of various challenging conditions, such as variations in pose [1], illumination [2], expression[3], occlusion [4], makeup spoofing [5]. All the above-mentioned issues occur while capturing the

K. R. Singh (✉) · R. S. Khedgaonkar
Department of Computer Technology, Yeshwatrao Chavan College of Engineering, Nagpur, Maharashtra, India
e-mail: ksingh@ycce.edu

R. S. Khedgaonkar
e-mail: rskhedgaonkar@ycce.edu

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_68

facial images. Because these changes in facial images are not permanent, they are classified as non-surgical FR issues. In addition to the concerns stated above in the field of face recognition, aging [6] and facial plastic surgery [7] have surfaced as new challenges for FR. FPS is divided into two categories: global plastic surgery (GPS) and local plastic surgery (LPS). In most cases, GPS is used to completely alter the facial structure after it has been functionally damaged. GPS may also be utilized to completely alter the appearance, texture, and other geometries of the face, making it difficult for any FR system to distinguish between prior and post-surgery process. Figure 1 depicts an example of GPS. Local plastic surgery (LPS), in contrast to GPS, is the procedure of altering particular face features to erase scars, blemishes, birthmarks, and damage caused due to an unfortunate incident. Rhinoplasty (nose surgery), Blepharoplasty (eyelid surgery), lip augmentation, cheek implant, Otoplasty, and Cranioplastyin are a few of the existing local plastic operations [7]. Figure 2 depicts celebrity facial images that have undergone LPS. Furthermore, the transformation of a facial feature alters an individual's facial appearance even more.

Fig. 1 Example of global plastic surgery

Fig. 2 Sample face images of celebrities undergone LPS. Top row: face image prior and post nose job. Bottom row: face image prior and post lip augmentation

This cause FR a tedious task that may change the appearance of the face permanently. Singh et al. was the first to investigate this unknown area of plastic surgery-based face recognition in a publication [7].

These challenging issues of huge variations in facial appearances of a person either because of non-surgical uncontrolled conditions or surgical conditions encourage researchers to understand the influence of each on face recognition. In this chapter, face recognition under non-surgical and surgical conditions has been performed using a patch-based approach [8]. The patch-based method divides the facial photos into patches initially. Furthermore, facial features are extracted for recognition employing a method for nearest matching using a classifier for each patch of the face under consideration. The performance of the patch-based method is evaluated on the IIIT-D facial databases [9] and the ORL face database [10]. The main contributions of the proposed chapter are: firstly, to understand the face recognition problem as surgical and non-surgical face recognition and impact of each, respectively. Secondly, to discuss the patch-based method to tackle the issues of non-surgical and surgical faces.

The organization of the proposed chapter would be as follows. The proposed work presents the working. Subsequently, the experimentation section focuses on the exhaustive experimentation performed on various surgical and non-surgical facial databases. Finally, the conclusion section is outlined followed by the scope for future work.

2 Proposed Work

For surgical and non-surgical face recognition, the chapter presents an approach called the patch-based approach [8]. In the first step of the patch-based approach, we split the facial image into three different patches. Then, we find the set of statistical features, and then these features are used to design a k-nearest neighbor classifier for the classification of surgical and non-surgical facial images into correct classification or misclassification responses.

An algorithm for patch-based face recognition is given in Algorithm 1.

Algorithm 1: Patch-based Approach
Input:
Train samples X= {$x1$, $x2$ … xn}, where n is the number of total training samples
Test samples Y = {$y1$, $y2$ … ym} where m is the number of total of test samples
Output: Predicted image {R}
Step 1: split train/test image into three horizontal patches as
x_i = {x_{i1}, $xi2$, … xij}, (three patches for i^{th} training sample, represented as x_{ij}, where i:1to n, j: 1 to 6)
y_i = {y_{i1}, y_{i2}, … y_{ij}}, (three patches for ith test sample, represented as y_{ij}, where i:1 to m, j: 1 to 6)
Step 2: compute statistical features for all x_i and y_i as F = {*mean; standard deviation; entropy; kurtosis; skewness; uniformity*}
Step 3: store all F for train images $x_i \epsilon X$ in look-up table for matching
Step 4: construct k-NN classifier for i = 1 to n do
compare all F_i in look-up table for X with all F of test image {yi}
Step 5: return nearest matched image {R}

3 Experimentation and Results

The patch-based method is tested on two different faces databases such as PSD [9] and ORL face database [10]. Since the main focus is on the surgical and non-surgical FR problem, precisely only frontal faces are chosen from the face databases. Each face database is divided in separate training sets and testing sets. In forthcoming section firstly, we present the experimentation and results for surgical face database followed by non-surgical face database.

3.1 Using Surgical Face Database

As discussed earlier, the three different types of facial plastic surgery, blepharoplasty (eyelid surgery), lip augmentation, and rhinoplasty (nose surgery) have been considered in the proposed work. The plastic surgery face dataset (PSD) [9] is comprised of

197 face image samples of blepharoplasty, 15 face samples of lip augmentation, and 329 face samples of rhinoplasty of pre-surgical and post-surgical facial images. Each facial image of an individual subject is resized to 273 × 238 pixels. Figure 3 depicts few tests of rhinoplasty, blepharoplasty, and lip augmentation from the PSD dataset [9]. For each 15 face samples undergone various surgeries from the top row are taken to form the training set. Similarly, 15 face samples undergone various surgeries from the bottom rows of Fig. 5 are selected to form the test set. In addition, each train and test image is subdivided into three patches of size 91 × 238 pixel size as shown in Fig. 4. Further, the statistical features of each patch are processed as mentioned earlier.

Consider the set of statistical parameters from given patches of train images is represented as $F = \{mean; standard\ deviation; entropy; kurtosis; skewness; uniformity\}$. The obtained features are thereafter used as an input to the KNN classifier for recognition. The experimentation has been performed for the different values of k (1, 3, 5, and 7) in order to find the best value of k. For different values of k, the

Fig. 3 Topmost row: training samples and bottommost row: testing samples from PSD database [9]. Row 1, Row 3, Row 5 shows the pre-surgical face images and Row 2, Row 4 and Row 6 shows the post-surgical face image for rhinoplasty, blepharoplasty, and lip augmentation, respectively [9]

Fig. 4 Patch-based partitioning

Fig. 5 Sample images from ORL face dataset with upright-left position in frontal view and left–right rotation [10]

recognition rate for blepharoplasty, lip augmentation, and rhinoplasty is tabulated in Table 1.

From Table 1, it can be observed that among all the three local facial plastic surgery under consideration, Rhinoplasty has outperformed on an average. Even under stricter condition with $k = 1$ Rhinoplasty has shown better results comparatively. The reason could be in Rhinoplasty the changes in bone structure of the nose does not affect the nearby region of the face, whereas this might not be the case Blepharoplasty and lip augmentation. Improvement in performance from $k = 1$ to higher values is easily noticable from the results tabulated in Table 1. From the results it is also evident that among all the three local facial plastic surgeries Blepharoplasty majorly effects the

Table 1 Comparison of the recognition rates in percentage (%) for three different surgeries on the PSD dataset using patch-based approach (different values of k for surgical face recognition)

Values of k	Types of LPS		
	Rhinoplasty (%)	Blepharoplasty (%)	Lip augmentation (%)
1	60	20	30
3	80	55	50
5	80	55	50
7	85	57	55

facial region which in turn results in poor performance. In next section of the chapter, we present the experimentation using non-surgical face dataset.

3.2 Using Non-surgical Face Dataset

The ORL face dataset is used as one of the non-surgical face databases. It comprised of 400 images of size 112 × 92 pixels. We used ten images of various people, each with 40 different subjects. Images of various subjects were captured at unusual times, with distinct facial emotions and lighting circumstances. The frontal faces acquisition is done in an upright position and with a left–right rotation. The experimentation is carried out on a frontal face image with an upright position and upleft rotation. Figure 5 presents few samples of the ORL face dataset.

As discussed earlier in surgical face datasets for different Image patches, statistical features are extracted for each patch. The features we have extracted are then inputted to the KNN classifier for recognition with different values of k. Table 2 shows the overall recognition rate of patch-based approach with different values of k as 1, 3, 5, and 7. From the obtained recognition rate we say that the value of k as 3, 5, and 7 gives a better result than the value of k as 1.

Table 2 Comparison of the recognition rates of ORL face database under patch-based approach under various values of k for non-surgical face recognition

Values of k	Upleft rotation (%)	Slight left–right rotation (%)	Upright rotation (%)
1	13	57	33
3	33	85	53
5	33	85	60
7	40	85	66

4 Conclusion

This chapter examined a patch-based technique for surgical and non-surgical facial datasets for all mentioned types of FPS in earlier sections in the context of surgical face recognition. ORL face dataset is considered with upright, upleft, and slight left–right rotation in the context of non-surgical face recognition. We used statistical characteristics and a KNN classifier to construct a patch-based technique. The PSD face dataset and the ORL face dataset were used in the experiments. Based on the findings, it is evident that the method is resistant to all the mentioned types of FPS. The presented approach gives an average performance of 63% and 48% for surgical face database and non-surgical face database, respectively, for values of k as 3, 5, 7. Based on the findings, it can be concluded that when surgical FR is performed, even though one local region of the face undergoes permanent change as a result of LPS, the other part of the face preserves its feature, potentially assisting in the recognition of the correct face and resulting in a greater recognition rate. Whereas in case of non-surgical face recognition the performance has been evaluated by considering three types of rotation viz; upright, upleft, slight left–right rotation where changes are non-permanent and face is completely positioned in the left–right direction which may greatly affect the performance of face recognition as compared to surgical face recognition.

One of the significant limitation in the proposed study is that we only tested a small number of test samples. The approach can be tested with fair size of test dataset, and also the results can be measured with other methods presented earlier. Furthermore, to improve results of surgical and non-surgical FR, one can employ other state-of the-art feature extraction approaches.

References

1. Liu X, Chen T (2005) Pose-robust face recognition using geometry assisted probabilistic modeling. In: Proceedings international conference on computer vision and pattern recognition. Washington, DC, USA
2. Li S, Chu R, Liao S, Zhang L (2007) Illumination invariant face recognition using near-infrared images. IEEE Trans Pattern Anal Mach Intell 29(4):627–639
3. Michel P, Kaliouby R (2003) Real time facial expression recognition in video using support vector machines. In: Proceedings of the 5th international conference on multimodal interfaces. USA
4. Priya NG, Wahida Banu RS (2014) Occlusion invariant face recognition using mean based weight matrix and support vector machine. Indian Acad Sci 7(2):154–160
5. Chen C, Dantcheva A, Swearingen T, Ross A (2017) Spoofing faces using makeup: an investigative study. In: IEEE international conference on identity, security and behavior analysis. New Delhi
6. Lanitis Taylor CJ, Cootes TF (2002) Toward automatic simulation of aging effects on face images. IEEE Trans Pattern Anal Mach Intell 24(4):442–450
7. Singh R, Vatsa M, Bhatt HS, Bharadwaj S, Noore A, Nooreyezdan SS (2010) Plastic surgery: a new dimension to face recognition. IEEE Trans Inf Forensics Secur 5(3):441–448

8. Khedgaonkar RE, Singh KR, Raghuwanshi MM (2018) Patch-based face recognition under plastic surgery. In: First international conference on secure cyber computing and communication (ICSCCC), pp 364–368
9. Indraprastha Institute of Information Technology Delhi (2008). Retrieved from http://www.face-rec.org/databases/psd
10. Cambridge University Engineering Department. AT & T (2018). Retrieved from http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att_faces.ta r.z.

RATSEL: A Game-Based Evaluating Tool for the Dyslexic Using Augmented Reality Gamification

D. Sorna Shanthi, Priya Vijay, S. Sam blesswin, S. Sahithya, and R. Sreevarshini

Abstract Dyslexia is a learning disorder that disrupts the ability of a person to understand and manipulate language during reading or writing. Children between the ages of 5 and 12 will show signs of the disorder as they begin school, where they will focus heavily on learning. The regular methodology of teaching and evaluation is not always supportive of kids with dyslexia. These kids need the support of the environment and society in order to improve accessibility to challenges and difficulties faced by them. ICT enables valuable support for dyslexia kids through the visual and auditory processes of learning. In this paper, we have proposed RATSEL, an Augmented Reality (AR) enabled game-based evaluation tool for enhancing the existing school examination methodology in order to sustainably support kids with dyslexia. RATSEL: the application holds a lot of subjective games which provide visual and auditory assistance in exercises and reading activities such that they are easier to grasp and understand. The main objective of RATSEL is to provide dyslexic kids with a game-based environment for evaluation using augmented reality technology, which transforms the evaluation processes in a more encouraging and hopeful manner.

Keywords ICT tool · Assessment · Augmented reality · Dyslexic · Basic education

1 Introduction

Children across the world face many types of cognitive diseases, and dyslexia is one of them. The traditional way of teaching is not always helpful for kids with dyslexia. Our current educational system is not sufficient to take into consideration

D. Sorna Shanthi (✉) · P. Vijay · S. Sam blesswin · S. Sahithya · R. Sreevarshini
Rajalakshmi Engineering College, Chennai, India
e-mail: Sornashanthi.d@rajalakshmi.edu.in

P. Vijay
e-mail: priya.vijay@rajalakshmi.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_69

exceptional cases where dyslexic kids cannot grasp concepts like other kids. Children with learning disabilities need to be educated using different approaches to achieve success in school. Providing proper awareness and accessibility to different tools and techniques can aid kids in integrating, managing, and completing basic education.

Augmented reality is one of the technologies in developing ICT-based interactive tools that can be used to enhance the interaction between kids, the real world and digital objects. Augmented reality (AR) is the overlaying of virtual information on the physical world. This typically means that the visual information is delivered through a handheld device like a tablet, phone or head-mounted display. Augmented reality (AR) is reliant on markers, or flashcards. AR technology detects the markers or flashcards and supplants the representation of the flashcard with an overlaid virtual object. AR flashcards can make learning entertaining for kids. Dyslexia is a lifelong condition, but it causes the most struggles throughout childhood and makes it extremely difficult to deal with basic education. Early diagnoses and assistance for the special needs of dyslexic children will prove very useful and life-changing. The proposed paper is based on this evidence. Dyslexic kids show signs of depression, low self-esteem, and also become unmotivated and develop a dislike for school. If this problem is left untreated, it can jeopardise their success in life. To overcome the difficulties faced by dyslexic kids, this paper proposes an AR-based application that would help the kids learn things at the same pace as a normal child.

The use of interactive multimedia applications result in increased reality in a fun, easy way will facilitate the pedagogic process for specialised children and help those young hearts to overcome their learning difficulties.

Smartphone and augmented reality technology were used to create a plethora of applications aimed specifically at educational purposes, as well as to teach and educate dyslexic children at various levels. ICT technologies increase the understanding level of students for principles and phenomena of subjects which require practical learning rather than just textual information, where learning from text-based knowledge is deficient. The use of augmented reality in such areas can greatly enhance the level of understanding and learning of kids for topics which require practical experimentation. AR can really come into play and demonstrate visuals in 2-D or 3-D forms.

Exhausted research has been done in the field of theoretical and practical aspects and benefits of augmented reality for interactive learning, but only few applications have been developed and used for real-time learning. Furthermore, practical researchers frequently focus on the implementation of augmented reality (AR) based on series games or gamification models for educational and teaching purposes. AR-based applications can be used everywhere, at any time by installing the application on smart phones.

The AR-based system enhances not only game-based learning but also augments the assessment of kids in a gratifying way. Assessment plays a vital role in testing whether the kids have learnt the skills taught. Games-based evaluation helps to

improve a positive attitude and interest towards taking assessment. Game-based evaluation helps to improve interest and time benefits rather than the traditional paper-based approach, since dyslexic students are nervous about taking up tests using pen and paper.

The application proposed in this paper aims to deliver a run-through and evaluation tool for dyslexic kids that can be used in schools and in educational institutes that specially train dyslexic kids by teachers to assess the progress of dyslexic kids and help them learn better in an interactive and pleasing manner. Kids themselves can use this application for practise purposes because of its usability and attractive human interface interaction through games.

2 Literature Survey

Hautala et al. [1] conducted a survey on the identification of reading difficulties in school children. The authors made an analysis and augmented the benefits of game-based learning over traditional approaches. The paper proved that screening and assessing reading skills is possible with diminutive game-based applications with an automatic rewarding system, even for students without learning disabilities. The research paved the way for the use of game-based learning and assessment for dyslexic children, who require a lot more interest in assessment than the average child.

In modern educational systems, according to Karamanoli and Tinakos [2], reading difficulties affect 3.5% of children. The AR application was designed to enhance the existing school book in order to help kids with dyslexia. The Aurasma platform was used to develop the application. It addresses children at the age of six, namely, it is for the first grade of primary school. The application, combined with a SQL server, can store the answers given by the kids to the book exercises. Teachers can have access to these answers and create their kids' portfolios for evaluating their progress. Also, with this application, kids can interact with their teacher by sending messages whilst they are at home. They gave a trail version to test and train primary school children.

Gupta et al. [3] developed the system improving accessibility for dyslexic impairments using augmented reality. In order to overcome the obstacles, the system employed augmented reality using a smartphone camera to enable the user to adjust the background text contrast ratio and text adaptation. After the application has been tested, the positive results are obtained by reducing their overall time for text passages by 21.03%. The application therefore proved to be a great help for the lives of people with dyslexia. The work inspired me to develop the proposed model to aid in the learning and assessment of preschool operations in dyslexic children.

Piyush et al. [4] demonstrated an instructive application for some individuals who pass by abnormal kids with learning disabilities without noticing in their paper augmented reality for abnormal kids. Dissimilar to different incapacities like viral fever, physical problems, and visual disability, a learning failure is a disguised

hindrance. The paper portrayed the usage of AR Scrapbook to teach kids in a fun, interactive, and compelling way. The paper outlines how to train cognitive skills and fine motor skills through AR.

Bhatti et al. [5] have worked to develop a cognitive learning framework for an interactive multimedia app that uses AR technology that is central to autism and allows children to interact with ART system. The system was designed to identify parameters of a dyslexic disorder. It included specially created marker images which could be printed using the standard mobile camera in a booklet format. The identified marker images were processed by the system and then overly 3-D animated objects will be added to the 3-D image on the screen. The purpose of this application was to enable students to memorise concept through AR and 3-D modelling but not student evaluation.

Rosemarie et al. [6] in the work showcased, augmented reality has a more immediate application because it requires no equipment beyond a smartphone. The paper integrated game elements into each component of the mobile game-based e-learning system for the basic reading of the Filipino language for dyslexic children.

Ibrahim et al. [7] showcased the game education architecture to create AR-based training and evaluation of dyslexia students by playing games through the app. The result was improved word and number pronunciation and spelling efficiency in dyslexic students. The work did not concentrate on shapes and other geometric structures in gamification.

The state of art clearly showcased the pathway and motivated to create an application to teach, test, and continuously train the basic education of dyslexic kids. The limitations identified by the state of work are

- The survey obviously exhibited that assessment and continuous monitoring of dyslexia kids play a vital role in improving the learning habits of the kids.
- Unlike virtual reality, however, there are not many educational apps in India to help dyslexic kids who find it difficult to read and write. The prevailing apps in our country focus only on video content that attracts kids.
- The applications that exist only visualise the 3-D model of the object scanned and do not help the kids to assess their disability in various aspects.

The main objective of the project is to create a virtual learning platform using augmented reality technology that would be a central tool that could benefit dyslexic students and their educators by transforming the assessment procedures in a more stimulating and entertaining manner. The objectives are subdivided as follows:

- Identifying subjects and levels of topics for dyslexia kids from 4 to 6 years old in basic education.
- Developing a test game-based model to motivate learning in an attractive and easy-to-use manner. A score
- Score rewarding system has been developed to update the status of the assessment for students.
- Integrating the system with a database so that continuous monitoring can be done by teachers, parents, or students

In the proposed work, an augmented reality-based system assessment tool is developed to assess the dyslexic kids.

3 System Framework and Methodologies

The RATSEL system framework is structured around developing subjective, fun-based interactive games for evaluating dyslexic kids using AR. In the first phase of RATSEL development, a survey and discussion were conducted to decide upon the subjective topics and evaluation methodology for developing fun-based interactive games to evaluate dyslexic kids. The second phase of this system involves user login into the Unity software and developing games using the AR Foundation platform API to support augmented reality. In the third phase, the AR application is integrated into the mobile platform to create the final deliverable app. Finally, in the fourth phase, the deliverable app will be field tested and assessed to collect feedback about the model so that corrective measures can be taken. In order to achieve the goal project, the following procedure is followed as shown in Fig. 1.

```
Step 1: Survey on Subjective Topics and Evaluation
Methodology to decide the game model.
Step 2: Flash card designing and 3D Model creation.
Step 3: Developing AR Games through Unity and AR
foundation.
Step 4: Integration of AR games with Mobile platform
Step 5: Deployment and launching of final AR
Application.
```

Fig. 1 Algorithm for RATSEL

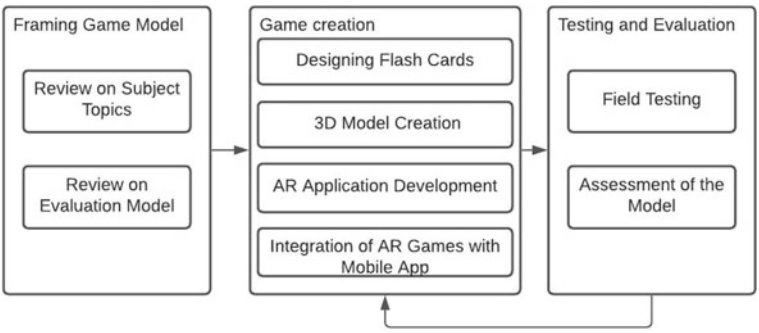


Fig. 2 Framework for RATSEL

Table 1 Subjects and levels identified for 4 to 6 years in basic education

Subject	Level 1	Level 2	Level 3	Level 4	Level 5
General	Identifying shapes	Identifying colours	Identifying animals	Identifying day/month	Identifying direction
Math	Identify numbers	Counting numbers	Matching number and objects	Addition	Subtraction
English	Alphabets	Words	Spelling check	Opposites (thin thick, small large, big and small …)	

The system framework is shown in Fig. 2. The framework depicts the overall view of the RATSEL game model, from framing topics to field testing and assessment of the accuracy of the model built.

The framework has three parts: framing the game model by identifying topics and its level; game creation, which involves the steps of creating a game-based test and training model; and finally testing and evaluation, in which the application is field tested as usability testing to find the accuracy and efficiency of the system.

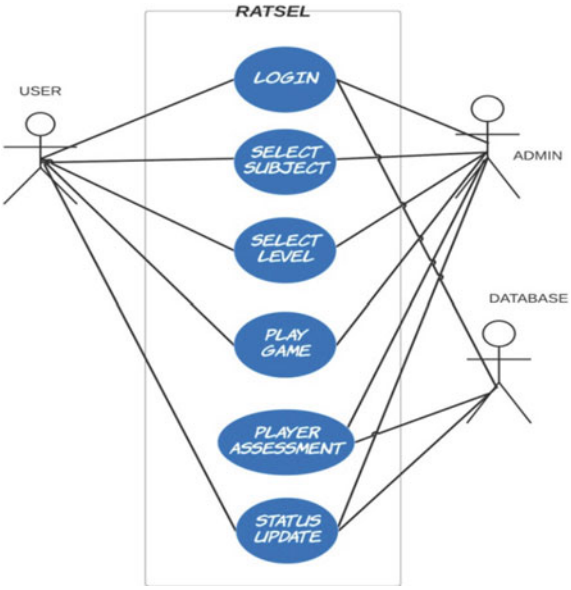
3.1 Framing Game Model

The research work was initiated by first collecting data and analysis of various aspects of augmented reality development for educational purposes. The literature survey was analysed, limitations were framed, and the subjects for the basic learning system for kids from 4 to 6 years were identified. Then topics were decided based on increasing complexity. Because all of the review papers focus on numbers and alphabets, the proposed work identified three major categories of subjects: general, which covers general knowledge and science, English, which includes alphabets and words, and math, which includes numbers, counting, subtraction, and addition in basic education. All three subjects are again categorised as five levels for training as well as assessment. Then a training and testing model was developed for the subjects and levels identified. The subjects and levels identified for 4 to 6 years in basic education are shown in Table 1.

3.2 Game Creation

The game creation step involves designing the flashcards, 3-D models, and development of the AR-based games through unity and the AR foundation. In the following sections, the modules for game creation are discussed in detail.

Fig. 3 Use case of RATSEL



Then a game model was prototyped to include train and test model as shown in use case Fig. 3.

AR Image Recognition

AR Image Recognition has been implemented in various games to train and evaluate dyslexic kids in the age group between 4 and 6 by recognising alphabets and shapes using 3-D visualisation of images. It also recognises objects that subsidy in engaging kids to focus on mathematical calculations like addition and subtraction using fun games. For example, in the alphabet category, level 1, an object will be displayed in the application so that the player (a dyslexic kid) will scan the flashcard that matches the object displayed. In the number category, addition and subtraction problems will be displayed on the screen, and the player should scan the flashcard for the right answer as shown in Fig. 5. Image tracking allows AR applications to detect, track, and augment 2-D images. The multiple target functionality enables the tracking of more than one image at a time.

AR Plane Detection

Dyslexic kids find it difficult to identify directions, identify colours, and differentiate between larger and smaller quantities. They also have difficulty counting numbers. To overcome these difficulties, AR Plane Detection has been implemented in games to make them understand these concepts in a better way. Plane detection finds horizontal and/or vertical planes in your physical surroundings, depending on the enabled detection mode (s) as shown in Fig. 4. Horizontal plane detection has been implemented in a few games, such as matching the colour of the car by choosing the right colour that has been displayed on the car. In the direction game, a monkey will be

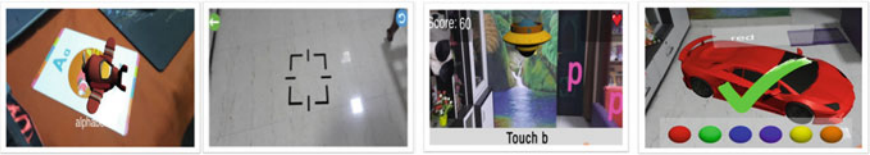


Fig. 4 Game creation development frames

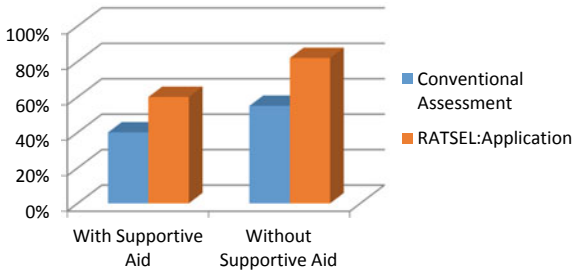


Fig. 5 Comparative study of conventional assessment and RATSEL

displayed and the player has to follow the instructions given by the monkey in order to complete the task according to the directions specified. In vertical plane detection, a shooting game has been implemented where the kids will have to shoot a particular target and learn about the months in a year.

AR Session

The AR Session is implemented on any game object as dyslexic kids find it very difficult to differentiate between alphabets such as B and D, P, and Q. For example, in the alphabets category of the game, an UFO will be moving around the screen spilling a mixed set of alphabets, and the player has to identify the correct alphabet that matches with the right alphabet that has been displayed on the instructions which is depicted in Fig. 5. AR Session manages the AR system's state and handles the session lifecycle. This class is the main entry point to the AR Core API. This class allows the user to create a session, configure it, start or stop it, and, most importantly, receive frames that allow access to the camera image and device pose. An AR scene should include an AR Session component. The AR Session controls the lifecycle of an AR experience, enabling or disabling AR on the target platform.

Voice Recognition

Voice recognition is implemented in the games for the kids to practise words and their spelling. To improve their reading ability, they read out loud the words that are displayed on the screen, which makes it easier for them to spell them out. As a result, their reading ability is being improved. Voice recognition is now being used in more complex games, which helps the players (dyslexic kids) to utilise speech recognition as a means to communicate and interact with the games. As a result, their reading ability is being improved.

RATSEL Login
In the first phase of the system, the user would be able to login with their credentials and have a separate database for their own so as to keep track of their progress, which would help them as well as their mentors to know where each individual stands. There is a level-based monitoring system that's been implemented so that the individual will not be able to move on to the next level unless he/she passes the current level. Microsoft Azure PlayFab is used to develop a login environment where the user would provide their credentials and login to the system, and the total score updated from all the categories of the game will be updated in the database when the player completes every level.

UX Design
UX experience is very important when it comes to dyslexia kids. In RATSEL, many buttons such as jump, shoot, select, drag, drop, place, target, and many more are used to kindle the kids' desire to attend the assessment as shown in Fig. 5.

3.3 Testing and Evaluation

In many respects, digital game-based assessment is very much better than conventional assessments. In addition, assessing the basic learning of dyslexic kids through an interactive game-based AR application is more effective than conventional approaches. RATSEL is such an application which encourages a stipulating game-based model for evaluating the basic learning of dyslexic kids. Whenever an individual completes every level, his/her scores will be updated in the database, and their level completion will also be assessed and monitored so that the teachers can help the kids with their particular disability, in which they are a little weak.

Field testing is a vital part of the RATSEL. In collaboration with AIKYA Learning Centre, the application was tested with 15 students. The test results proved that the game-based assessment methodology using AR was more effective than conventional assessment. The comparative study of the different assessment results is shown in Fig. 5.

The application was also tested against the User Experience. Users and the Supportive teachers enjoyed the assessment model and provided the feedback on the application based one ease of use, attractiveness and Realism of assessment. Figure 6 shows that the feedback was 80% good on a positive note.

4 Conclusion

Research says that children with dyslexia seem to have difficulty with coloured overlays, differentiation with certain alphabets, reading disabilities, direction identification, and counting. Thus, there is a pressing need for further testing and research

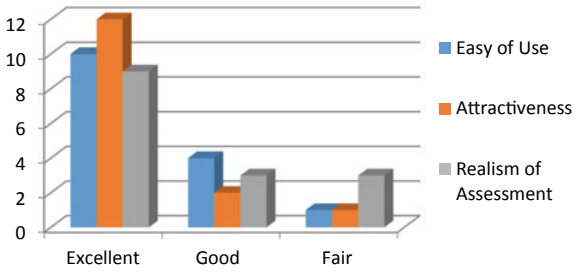


Fig. 6 Feedback of users experience on RATSEL

to be conducted to overcome these disabilities. This application serves as a tool that not only helps the dyslexic kids by providing practise but also keeps a track of the progress of each individual that might help the teachers to understand the kids better and focus on their difficulties. The application's evaluation from teachers showed that there is a positive attitude towards the incorporation of AR in dyslexic kids' education. Further research may be focused on enhancing the application according to AR development, and we look forward to increasing the number of levels and genres of games with respect to the growing syllabus. Small steps have been made towards the application of AR in dyslexics' education, but these steps can be the foundation of tomorrow's educational system, where learning disabilities, like dyslexia, will be treated with respect and attention.

References

1. Hautala J, Heikkilä R, Nieminen L, Rantanen V, Latvala J-M, Richardson U (2020) Identification of reading difficulties by a digital game-based assessment technology. J Educ Comput Res 58(5):1003–1028. https://doi.org/10.1177/0735633120905309
2. Karamanoli P, Tsinakos A (2016) Augmented reality and dyslexia: a new approach in teaching students. In: International conference on blended learning for the 21stcentury learner
3. Gupta T, Sisodia M, Fazulbhoy S, Raju M, Agrawal J (2019) Improving accessibility for dyslexic impairments using augmented reality. In: 2019 international conference on computer communication and informatics (ICCCI), pp 1–4. https://doi.org/10.1109/ICCCI.2019.8822152
4. Piyush D, Navin KS, Himadri B, Aryan G, Vairamuthu S (2019) Augmented reality for abnormal kids. Int J Sci Technol Res 8(11)
5. Bhatti Z, Bibi M, Shabbir N (2020) Augmented reality based multimedia learning for dyslexic children. In: 3rd international conference on computing, mathematics and engineering technologies, pp 1–7. https://doi.org/10.1109/iCoMET48670.2020.9073879
6. Rosemarie T, Arispe MCA, Torio JO, Maligat Jr DE (2020) Mobile game-based learning to enhance the reading performance of dyslexic children. Int J Adv Trends Comput Sci Eng 9(1.3):332–337
7. Ibrahim A, AzaJaiza M, Farah Farhana R, Norziah S (2020) DyslexiAR: augmented reality game based learning on reading, spelling and numbers for dyslexia user's. In: 7th international conference on computational science and technology, LNEE, vol 724. Springer, pp 259–269

Design of Mini-Games for Assessing Computational Thinking

V. V. Vinu Varghese and V. G. Renumol

Abstract Computational Thinking (CT) is a process that allows us to solve problems in an efficient and effective manner using a computer. Many of the educational institutions around the world have adopted CT into their curriculum because it helps to foster the core competencies of the twenty-first century. There are different methods for imparting and improving CT but are not limited to interventions, scaffoldings, games, AR/VR, robotics and unplugged activities. Games and Gamification are interesting domains for educational researchers. Gamified approaches, including serious and mini-games, provide a platform that enables educators to impart, foster and assess CT skills while maintaining the interest, motivation and fun element of games. This research paper proposes a design of an interactive tool for assessing CT. The proposed tool allows the assessment of CT cornerstones in a fun and motivating environment that alleviates the stress and boredom that may occur while attending the traditional multi-choice assessment.

Keywords Game design · Computational thinking · Automatic skill assessment · Mini-games · Stealth assessment

1 Introduction

Computational Thinking (CT) is a skill that enables a person to solve a computational problem effectively. In 2006, Wing [1] introduced CT as an approach to solve problems, building systems, and understanding human behavior that draws the power and limitations of computing. Over the years, many have developed different interpretations of CT. However, in 2010 Wing introduced a more refined definition of CT, stating, "CT is a thinking process in which solutions are represented in a form that can be effectively carried out by an information-processing agent" [2].

In 2012, the Royal Society defined CT as "Computational thinking is the process of recognizing aspects of computation in the world that surrounds us, and applying tools

V. V. Vinu Varghese (✉) · V. G. Renumol
Cochin University of Science and Technology, Cochin, Kerala, India
e-mail: vinghese@gmail.com

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_70

and techniques from Computer Science to understand and reason about both natural and artificial systems and processes" [3]. Even though there is a lack of consensus on the formal definition of CT, it did not affect the research in this domain as the definition of CT may vary based on the research conducted [4]. CT is intertwined with computer science and is closely related to programming; however, CT is not the same as programming. Many of the researchers have incorporated programming elements in CT-based research [5–8].

As per the literature reviewed, it is observed that CT has two major research areas—development of CT and assessment of CT [9]. The former consists of imparting and fostering CT in students ranging from kindergarten to graduate students. The latter one is the research based on the assessment of CT skills in students. A few of the researchers have studied automatic/real-time assessment of CT skills [10–13]. The abovementioned researchers used real-time assessment and educational data analytics to predict the student's performance. Shute and Ventura [14], on the other hand, researched automatic/stealth assessment of students' skills and states that "traditional classroom assessment are detached events, rarely influence learning and have validity issues." He also identified that stealth assessment could reduce test anxiety without compromising the validity and reliability of the test [15]. Since only a few researchers have ventured into the domain of real-time, stealth and analytics-driven assessment of skills related to CT skills and the widespread consensus that the use of games in education provides motivation and engagement [16–18]. We intend to utilize this gap to contribute to the research community by proposing a tool for automatically assessing CT skills using serious mini-games as a medium that can alleviate the stress of participants during traditional assessment methods. The rest of the paper is organized as follows: Sect. 2 describes related work, Sect. 3 explains the proposed design and Sect. 4 concludes the paper.

2 Related Works

This section specifies the research done specifically in the domain of CT, such as CT games and CT Assessment.

2.1 Computational Thinking Games

Video games used for educational purposes are classified as serious games and mini-games. The former has only one objective/scenario in mind and is used for teaching a specific concept [19–21]. They are played only once or twice as per the requirement. Whereas mini-games, shorter versions of the game, teach only a single concept [22]. They are easy to play and have straightforward rules. Video games are widely used to foster CT [23]. AI games and puzzles were used to impart CT concepts in engineering students [24] and this gamification approach increased the confidence in students to

solve problems using CT. The work done by Jiang et al. [25] provided a better understanding of game design concepts for serious games. Using the game Gidget [26], a debugging first approach was proposed and researched to foster computing education in teenage students. Significant improvement of computing skills was reported in this study.

2.2 Computational Thinking Assessment

There is a broader area of research that consists of a variety of tools used to assess CT. Based on a literature review [9], the research on CT assessment is classified as questionnaire-based and software-based assessment tools. The questionnaire-based tools include but are not limited to the Computational Thinking Test (CTT) [27], which utilizes a visual answering system to assess the CT skills. Bebras [28] is another challenge in which visual elements are used to teach informatics and assess CT skills simultaneously. CS-Unplugged activities [29] assess students' intrinsic motivation in an unplugged activity setting. Some of the software-based tools are Fairy Assessment [30], which utilizes the block-based programming environment in Alice to measure the CT skills; Computational Thinking Pattern (CTP) [31], uses the AgentSheets environment to assess whether students are able to transfer the knowledge gained from game creation assignment to create science simulations and Dr. Scratch [32] that uses projects done in Scratch programming tool to access the CT skills.

3 Proposed Design

This section showcases the high-level architecture of the proposed tool, the design of each mini-game along with the wireframes, the assessment of CT skills, game mechanics, game assessment methods and target population.

3.1 Architecture

This paper proposes an interactive tool where the participants' CT skills are measured while playing mini-games. The high-level architecture of the proposed tool is shown in Fig. 1. The tool consists of five components. The first component is the user who is able to use the tool. The second is the User Interface component which is responsible for the interaction with the user. The third component is the Game Engine. It is the most critical part of the system, which is responsible for managing the user interface and the Assessment Engine. The Assessment Engine is the fourth component. It is

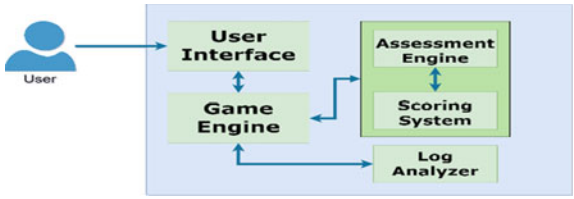


Fig. 1 Block diagram

responsible for the real-time assessment of the users. The Scoring System is a sub-component of the assessment engine and is responsible for recording the score for each game played by the user and stores the final score. The fifth component is the Log Analyzer which is responsible for recording and storing each user's action in the game. These user logs can be used as input for educational data mining.

3.2 Game Design

We have developed multiple mini-games: Coffee Maker, Shortest Path, Jigsaw puzzle, Sorting, Electricity calculator, Wash the clothes, Pacman, Plumbing puzzle, Maze runner, Decode the string, Water puzzle and Konigsberg bridges puzzle. The games have various difficulties ranging from easy, medium and hard.

- The Coffee Maker game is associated with abstraction and decomposition. In this game, the participant is asked to categorize the items required to make coffee into three categories: Groceries, Utensils and Appliances.
- The second game is the Jigsaw puzzle which maps to Pattern Recognition, abstraction and algorithmic thinking. It asks the participant to create a complete picture by matching the pattern template or the picture provided.
- The third game is Shortest Path Maps to decomposition, abstraction and algorithmic thinking. This game is not related to graph theory but rather a simple problem-solving one. The participant is asked to find the shortest path from a given source to a destination**.** Each path is given a distance value, and the path with the lowest distance is the shortest one.
- The fourth game in the series is the Sorting game. It maps to Algorithmic Thinking, abstraction and pattern recognition. It asks the participant to sort a set of eight numbers in ascending order. The participant can swap a number to an adjacent location by using the temporary location provided.
- The fifth game is Electricity calculator with maps to abstraction and pattern recognition. The students are asked to switch on the circuits to meet the specific voltage. Turning the switches on and off displays the corresponding voltages represented by the switches in the display panel. The game is of hard difficulty.
- The sixth game is Wash the cloths and maps to abstraction and decomposition. The students are asked to set the washing machine to the specified time and mode. The game's difficulty is hard.

- The seventh game is packman with non-movable ghosts. The students are asked to reach the destination by collecting the berries and at the same time avoiding the ghosts. The game maps to abstraction and algorithmic thinking. The game is of easy difficulty and can be upgraded to medium and hard difficulty settings by removing the backtracking option from the challenge and providing movable ghosts.
- The eights game is a plumbing puzzle and maps to abstraction and decomposition. The students need to rearrange/rotate the pipes to allow water to flow from source to destination. The game is of medium difficulty.
- The ninth game is maze runner, which maps to abstraction and decomposition. The students are asked to identify the paths from source to destination. The game is of easy difficulty by integrating the shortest path algorithm along with pathfinding.
- The tenth game is Decode the string. The game provides a simple substitution cipher and asks the students to decode a particular string and maps to abstraction, pattern recognition and algorithmic thinking. The game is of medium difficulty and can be updated to the hard difficulty by changing the cipher algorithm.
- The eleventh game is a water puzzle and maps to abstraction and algorithmic thinking. The game consists of different flasks interconnected with each other. The students are asked to find out which flask fills the water first. The game is of easy difficulty.
- The twelfth game is the Konigsberg bridges puzzle that maps to algorithmic thinking and pattern recognition. The game consists of 7 bridges across two streams in a city. The students need to cross the bridge exactly once from the beginning of the journey to the end. The students are also allowed to add a new bridge across the stream anywhere to solve the problem. The game is of medium difficulty.

We have developed a wireframe for these mini-games. The tool is incomplete as of now. We are in the process of designing the user interface for these games. Unity 2-D will be used for the development phase. The wireframes for the user interface are listed as Figs. 2, 3 and 4.

3.3 Assessment of CT Skills

The proposed tool assesses four concepts/cornerstones of CT. They are Abstraction [33], Algorithmic Thinking [34], Decomposition [35] and Pattern Recognition [36]. These four cornerstones were identified and well defined in the literature [37] and have consensus among researchers. This design proposes a set of games under each cornerstone to assess the corresponding CT concept. The games are selected such that they can be solved by using a mathematical, logical or real-world concept. The mapping of CT cornerstones to games is listed under Table 1.

In the research conducted by Brennan and Resnick [36], CT was divided into Computational concepts, Computational practices and Computational perspectives.

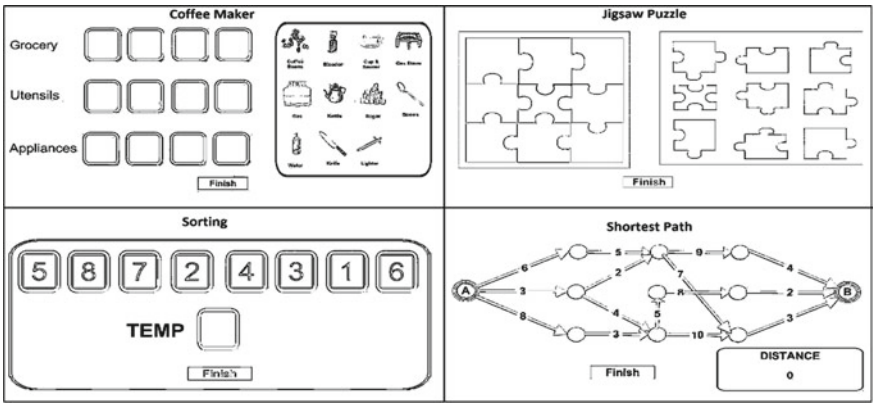


Fig. 2 Wireframe for coffee maker, jigsaw, sorting and shortest path game

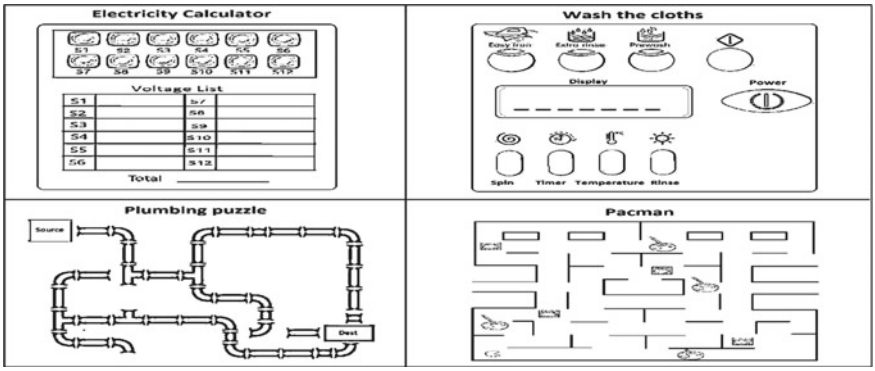


Fig. 3 Wireframe for electricity calculator, wash the cloths, plumbing puzzle and Pacman

These computational concepts are mapped into our proposed design. Out of these seven computational concepts, namely, sequences, loops, events, conditionals, operators, data and parallelism, the concept parallelism is excluded as it could not be directly incorporated into the game mechanics. Our design supports computational practices such as iteration, testing, reusing, remixing and abstraction. The mapping of computational concepts to games is shown in Table 2.

3.4 Game Mechanics, Assessment and Target Audience

This section intends to discuss our proposal for the gameplay mechanics, game assessment and target audience. The game mechanics specifies how the participants can play the game. The assessment section specifies the rubric related to the scoring

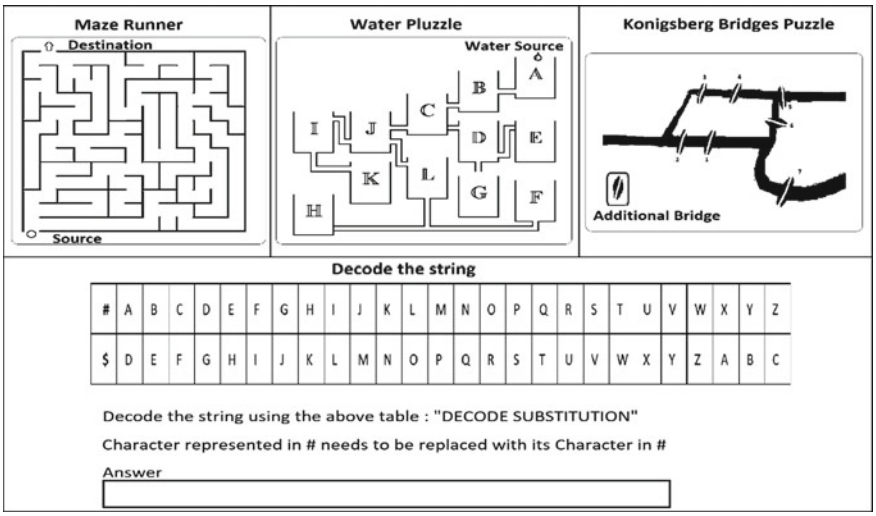


Fig. 4 Wireframe for maze runner, water puzzle, Konigsberg bridge and decode the string

Table 1 Mapping CT cornerstones to the games

CT cornerstones / Games	Abstraction	Decomposition	Pattern recognition	Algorithmic thinking
Coffee maker	x	x	–	–
Sorting	x	–	x	x
Shortest path	x	x	–	x
Jigsaw puzzle	x	–	x	x
Electricity calculator	x	–	x	–
Wash the clothes	x	x	–	–
Pacman	x	–	–	x
Plumbing puzzle	x	x	–	x
Maze runner	x	x	–	–
Decode the string	–	–	x	x
Water puzzle	x	–	–	x
Konigsberg bridges	–	–	x	x

and the target audience specifies the age group of the target population, why they were selected.

Game Mechanics: The game mechanics were selected with ease of use in mind. It includes drag and drop, single selection and repeated selection. The mechanics are kept relatively simple to reduce any discomfort to the participants and provide

Table 2 Mapping computational concepts to games

CT concepts / Game	Sequences	Loops	Events	Conditionals	Operators	Data
Coffee maker	–	x	x	x	–	–
Sorting	x	x	x	x	–	x
Shortest path	x	–	x	x	x	x
Jigsaw puzzle	–	–	x	x	–	–
Electricity calculator	x	–	x	x	x	x
Wash the clothes	x	–	x	–	x	x
Pacman	–	–	–	x	–	–
Plumbing puzzle	x	x	x	–	–	–
Maze runner	x	–	x	–	–	–
Decode the string	–	–	x	x	–	x
Water puzzle	x	–	–	x	–	–
Konigsberg bridges	x	x	–	x	–	–

feedback on the scenario. The games are categorized based on the game mechanics are as follows:

- *Drag and Drop*: We plan to include games such as Coffee Maker, Sorting, Jigsaw puzzle, Decode the string and Konigsberg puzzle to use drag and drop mechanics. The challenges in the abovementioned game require the user to move/rearrange the elements in-game.
- *Single Selection*: The games in this category require the user to select a single option from a list of options to solve the problem. When the user clicks a game element, it is either activated or deactivated. So, the games such as Shortest Path, Electricity Calculator, Pacman, Plumbing puzzle, Maze runner and Water puzzle are grouped under the single selection game mechanics.
- *Repeated Selection*: The game Wash the clothes utilize the repeated selection mechanic since the game elements/button is programmed for multiple actions as opposed to the single selection. These actions include setting a variable to a particular value via repeated press/selection.

Game Assessment: We propose a full score system in this paper. When a participant plays a mini-game and the target is achieved, a full score of 5 points is given to the CT cornerstones mapped to the particular game. The rubric proposed in this paper is listed in Table 3. The columns filled as "0" represent the CT cornerstone that is not evaluated as part of the assessment even though that particular skill is used in the mini-game. The mapped CT cornerstones are not considered in the rubrics to standardize the scoring system and round it to a total of 100 points. After the participant finishes all the games on the list, the student shall be provided with a detailed report on the points gained in each CT cornerstone.

Table 3 Rubric for the game

CT cornerstones score / Games	Abstraction	Decomposition	Pattern recognition	Algorithmic thinking
Coffee maker	5	5	–	–
Sorting	0	–	5	5
Shortest path	5	5	–	0
Jigsaw puzzle	0	–	5	0
Electricity calculator	5	–	5	–
Wash the clothes	0	5	–	–
Pacman	5	–	–	5
Plumbing puzzle	0	5	–	0
Maze runner	0	5	–	–
Decode the string	–	–	5	5
Water puzzle	5	–	–	5
Konigsberg bridges	–	–	5	5
Total score	**25**	**25**	**25**	**25**
Grand total	**100**			

Target Audience: The proposed mini-games are designed for participants of age groups ranging from 10 to 16 years. In this proposal, the age selection is based on Piaget's Formal operational stage [38], where the participants fully develop the CT concepts required to solve the game. The target audience may change after an expert evaluation study is conducted. The mini-games are designed in such a way that students with minimal or no programming knowledge can attend the assessment.

4 Conclusion and Future Work

In this paper, we have presented a set of mini-games for assessing the four cornerstones of Computational Thinking. The software is in the initial stage of its development and currently provides only limited functionalities. One of the features that distinguish our tool from similar assessment tools is that it allows students to solve problems through immersive games. The software is designed in such a way that the assessment of their CT skill is done simultaneously in real-time without the knowledge of the participant. We believe that this system may reduce the stress and boredom while attending traditional multi-choice assessment questions but need to be proved statistically. Further, we plan to develop a scale for CT after carefully

studying the usability, self-efficacy and validation factors. We also plan to do an empirical evaluation of our tool to support our claim.

References

1. Wing JM (2006) Computational thinking. Commun ACM 49:33–35. https://doi.org/10.1145/1118178.1118215
2. Wing JM Computational thinking: what and why? http://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why (2010)
3. The British Royal Society (2012) Shut down or restart? Br J Educ Technol 789–801
4. Berland M, Wilensky U (2015) Comparing virtual and physical robotics environments for supporting complex systems and computational thinking. J Sci Educ Technol 24:628–647. https://doi.org/10.1007/s10956-015-9552-x
5. Winter V, Friend M, Matthews M, Love B, Vasireddy S (2019) Using visualization to reduce the cognitive load of threshold concepts in computer programming. In: Proceedings—frontiers in education conference, FIE, pp 1–9. https://doi.org/10.1109/FIE43999.2019.9028612
6. Taylor K (2013) Can utilizing social media and visual programming increase retention of minorities in programming classes? In: Proceedings—frontiers in education conference, FIE, pp 1046–1048. https://doi.org/10.1109/FIE.2013.6684986
7. More A, Kumar J, Renumol VG (2011) Web based programming assistance tool for novices. In: Proceedings—IEEE international conference on technology for education, T4E 2011, pp 270–273. https://doi.org/10.1109/T4E.2011.55
8. Philip M, Renumol VG, Gopeekrishnan R (2013) A pragmatic approach to develop computational thinking skills in novices in computing education. In: Proceedings of the 2013 IEEE international conference in MOOC, innovation and technology in education, MITE 2013, pp 199–204. https://doi.org/10.1109/MITE.2013.6756334
9. Vinu Varghese VV, Renumol VG (2021) Assessment methods and interventions to develop computational thinking—a literature review. In: 2021 international conference on innovative trends in information technology, ICITIIT 2021. https://doi.org/10.1109/ICITIIT51526.2021.9399606
10. Koh KH, Basawapatna A, Nickerson H, Repenning A (2014) Real time assessment of computational thinking. In: Proceedings of IEEE symposium on visual languages and human-centric computing, VL/HCC, pp 49–52. https://doi.org/10.1109/VLHCC.2014.6883021
11. Moreno-León J, Harteveld C, Román-González M, Robles G (2017) On the automatic assessment of computational thinking skills: a comparison with human experts. In: Conference on human factors in computing systems—proceedings. Part F1276, pp 2788–2795. https://doi.org/10.1145/3027063.3053216
12. Grover S, Basu S, Bienkowski M, Eagle M, Diana N, Stamper J (2017) A framework for using hypothesis-driven approaches to support data-driven learning analytics in measuring computational thinking in block-based programming environments. ACM Trans Comput Educ 17:1–25. https://doi.org/10.1145/3105910
13. Eguíluz A, Guenaga M, Garaizar P, Olivares-Rodríguez C (2020) Exploring the progression of early programmers in a set of computational thinking challenges via clickstream analysis. IEEE Trans Emerg Top Comput 8:256–261. https://doi.org/10.1109/TETC.2017.2768550
14. Sell C, Baserga R, Rubin R (1995) Insulin-like growth factor I (IGF-I) and the IGF-I receptor prevent etoposide-induced apoptosis. Can Res 55:303–306
15. Shute V, Ventura M (2013) Stealth assessment: measuring and supporting learning in video games Valerie
16. Yue WS, Wan WL (2016) The effectiveness of digital game for introductory programming concepts. In: 2015 10th international conference for internet technology and secured transactions, ICITST 2015, pp 421–425. https://doi.org/10.1109/ICITST.2015.7412134

17. Wang T, Towey D (2013) A mobile virtual environment game approach for improving student learning performance in integrated science classes in Hong Kong international schools. In: Proceedings of 2013 IEEE international conference on teaching, assessment and learning for engineering, TALE 2013, pp 386–388. https://doi.org/10.1109/TALE.2013.6654467
18. Shabalina O, Malliarakis C, Tomos F, Mozelius P (2017) Game-based learning for learning to program: from learning through play to learning through game development. In: Proceedings of the 11th European conference on games based learning, ECGBL 2017, pp 571–576
19. Devisch O, Gugerell K, Diephuis J, Constantinescu T, Ampatzidou C, Jauschneg M (2017) Mini is beautiful. Playing serious mini-games to facilitate collective learning on complex urban processes. Interact Des Archit 141–157
20. Wiemeyer RDGE (2016) Serious games. Springer International Publishing, Cham. https://doi.org/10.1111/j.1740-9713.2016.00978.x
21. Crookall D (2010) Serious games, debriefing, and simulation/gaming as a discipline. http://sg.sagepub.com. https://doi.org/10.1177/1046878110390784
22. De Jans S, Van Geit K, Cauberghe V, Hudders L, De Veirman M (2017) Using games to raise awareness: how to co-design serious mini-games? Comput Educ 110:77–87. https://doi.org/10.1016/j.compedu.2017.03.009
23. Webb HC, Rosson MB (2013) Using scaffolded examples to teach computational thinking concepts. In: SIGCSE 2013—Proceedings of the 44th ACM technical symposium on computer science education, pp 95–100. https://doi.org/10.1145/2445196.2445227
24. Silapachote P, Srisuphab A (2016) Teaching and learning computational thinking through solving problems in artificial intelligence: on designing introductory engineering and computing courses. In: Proceedings of 2016 IEEE international conference on teaching, assessment and learning for engineering, TALE 2016, pp 50–54. https://doi.org/10.1109/TALE.2016.7851769
25. Jiang X, Harteveld C, Huang X, Fung AYH (2019) The computational puzzle design framework 1–11. https://doi.org/10.1145/3337722.3337768
26. Lee MJ, Bahmani F, Kwan I, Laferte J, Charters P, Horvath A, Luor F, Cao J, Law C, Beswetherick M, Long S, Burnett M, Ko AJ (2014) Principles of a debugging-first puzzle game for computing education. In: Proceedings of IEEE symposium on visual languages and human-centric computing, VL/HCC, pp 57–64. https://doi.org/10.1109/VLHCC.2014.6883023
27. Román-González M, Pérez-González JC, Moreno-León J, Robles G (2018) Can computational talent be detected? Predictive validity of the computational thinking test. Int J Child Comput Interact 18:47–58. https://doi.org/10.1016/j.ijcci.2018.06.004
28. Dagiene V, Stupuriene G (2016) Bebras—a sustainable community building model for the concept based learning of informatics and computational thinking. Inf Educ 15:25–44. https://doi.org/10.15388/infedu.2016.02
29. Jiang S, Wong GKW (2017) Assessing primary school students' intrinsic motivation of computational thinking. In: Proceedings of 2017 IEEE international conference on teaching, assessment and learning for engineering, TALE 2017, pp 469–474. https://doi.org/10.1109/TALE.2017.8252381
30. Werner L, Denner J, Campe S, Kawamoto DC (2012) The fairy performance assessment: measuring computational thinking in middle school. In: SIGCSE'12—proceedings of the 43rd ACM technical symposium on computer science education, pp 215–220. https://doi.org/10.1145/2157136.2157200
31. Basawapatna A, Koh KH, Repenning A, Webb DC, Marshall KS (2011) Recognizing computational thinking patterns. In: SIGCSE'11—proceedings of the 42nd ACM technical symposium on computer science education, pp 245–250. https://doi.org/10.1145/1953163.1953241
32. Troiano GM, Snodgrass S, Argimak E, Robles G, Smith G, Cassidy M, Tucker-Raymond E, Puttick G, Harteveld C (2019) Is my game ok Dr. scratch?: exploring programming and computational thinking development via metrics in student-designed serious games for STEM. In: Proceedings of the 18th ACM international conference on interaction design and children, IDC 2019. Association for Computing Machinery, Inc., pp 208–219. https://doi.org/10.1145/3311927.3323152

33. Dams D, Grumberg O (2018) Abstraction and abstraction refinement. In: Handbook of model checking. Springer International Publishing, pp 385–419. https://doi.org/10.1007/978-3-319-10575-8_13
34. Futschek G (2006) Algorithmic thinking: the key for understanding computer science. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics). 4226 LNCS, 159–168. https://doi.org/10.1007/11915355_15.
35. Barr V, Stephenson C (2011) Bringing computational thinking to K-12: what is involved and what is the role of the computer science education community? ACM Inroads 2:48–54. https://doi.org/10.1145/1929887.1929905
36. Brennan K, Resnick M (2012) New frameworks for studying and assessing the development of computational thinking. Annual American Educational Research Association meeting, Vancouver, BC, Canada, pp 1–25
37. Selby C, Woollard J (2013) Computational thinking : the developing definition. In: ITiCSE conference 2013, pp 5–8
38. Wadsworth B (1984) Piaget's theory of cognitive and affective development, 3rd edn.

Vehicle Cluster Development

Sanskruti Raut, Shravani Naware, and Vanita Tank

Abstract As the evolution of the technology in the automobile industry is taking place rapidly in recent times, this insists upon the need of developing an affordable and durable digital dashboard. This paper presents the work pertaining to the project which aims at developing a digital dashboard for automobiles using the model-in-loop, software-in-loop, and hardware-in loop methodology. In an automobile, vehicle cluster or a dashboard is the panel which has gauges featuring a speedometer, a tachometer, a fuel indicator, etc. Dashboards are of two types, namely analog dials and digital dials. An instrument panel cluster (IPC), or a vehicle dashboard, is one of the few devices in an automobile that communicates with the driver and since it is one of the few modules the driver can see it becomes an important aspect for the driver to get relevant information about the information and the status of the vehicle. This paper describes the design and implementation of a digital dashboard for an automobile, using MATLAB's Simulink for simulations. MATLAB's App Designer for designing a GUI and an onboard diagnostics port (OBD-II port) present inside a car and graphic programmable LCD. The LCD is programmed in a way so as to exhibit speed, engine RPM along with others, which are necessary parameters to be displayed to the driver while driving. The final device is low cost and ready to be adapted to other vehicles with low physical or software modification.

Keywords Automotive · MATLAB GUI · Simulink · OBD-II · Arduino · LCD

1 Introduction

The new industrial paradigm makes it facile for the common people to acquire an automobile. Since then, the automobile industry has grown exponentially and has

S. Raut (✉) · S. Naware · V. Tank
MIT World Peace University, Pune, India
e-mail: sanskrutiraut14@gmail.com

V. Tank
e-mail: vanita.tank@mitwpu.edu.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_71

shown tremendous development and increasing complexity as the needs and demands of the people have increased day by day. The advancements in technology make it a need to upgrade traditional systems into cost-efficient solutions to accelerate digital transformation. The instrument cluster or the gauge cluster plays an important role while driving a car. There are different displays and indicators that enable the driver to measure the parameters of a car such as speed, fuel indicator, and brake indicator. The function of the instrument cluster is to keep the driver informed about the current driving status of the car. The instrument cluster is mounted in the dashboard, just behind the steering wheel. The indicators that we have used include a speedometer, tachometer, fuel indicator, brake indicator, gear indicator, and check engine light. The traditional analog gauges work flawlessly when it comes to tracking the information as there is a physical movement of the needle which catches the eye of the driver while driving; however, the precision of the analog gauges may hamper due to parallax errors. As the recent trend, in the automobile industry, is about making the vehicles fuel efficient and smart, development on using microprocessors and electronic devices used onboard has increased. Not only does that aid the driver but also makes decisions when it comes to safety and security of the passengers. Due to their ability to make such complicated decisions, they are termed as electrical control units or ECUs, and they communicate with the entire vehicular network via protocols called CAN protocols. This communication happens through transmission of data packets containing real-time information from sensors or any error that has been detected. This information can be obtained through a port, present in all cars called on-board diagnosis or OBD port. The OBD-II adapter, ELM327 Bluetooth adapter, is used to fetch the real-time information from the car. The OBD-II adapter is connected to the port which is underneath the steering wheel in most cars and sets up a connection with the microcontroller that is the Arduino UNO. A Bluetooth module HC-05, which can communicate in two ways, that is, it is full-duplex and is interfaced with Arduino UNO. Its serial communication makes it easier to communicate and interface with the PC or the controller. After the connection is set up, the real-time information is obtained from the car and is fetched to MATLAB and further to MATLAB's App Designer to display the information on the dashboard created using MATLAB's GUI.

2 Literature Survey

The methodology followed for the research work is the MIL, SIL, and the HIL methodology. This paper [1] demonstrates the techniques to develop and build a model effectively. Through the model breakdown in terms of functional and interface layers and into three levels, it becomes easier to move through different developmental stages that is MIL, SIL, and HIL without having to maintain the multiple variants of the model. The research work pertaining to the project has been conducted using the same methodology. The model-in-loop, software-in-loop, and hardware-in-loop solutions are model-based design approaches that help to authenticate the

developed solutions with simulations of the real-time situations first, followed by the implementation of the code before deploying the hardware. This idea of following a modular approach helps in solving problems in a stepwise approach. At the initial MIL stage, the logic is tested on simulation platforms like MathWorks Simulink. At the SIL stage, the model is compiled into a C code and run with a fixed time step, and the nature of the components in the system remains the same as MIL. At the HIL stage, the model is interfaced using hardware, and the compiled code is deployed on the hardware.

Another paper [2] here discusses the design and implementation of a CAN-based digital dashboard unit for an electric car called SCE. The authors describe the importance of selecting particular color palettes suitable for the driver. Taking into consideration the direct sunlight, for highest visibility, a high-contrast color palette was selected. The paper describes the designing aspect of the dashboard along with the CAN bus implementation for the electric car. The CAN bus is used for serial communication between the vehicle's main control unit and the LCD via serial interface.

Authors of this paper [3] demonstrate the design and implementation of CAN bus protocol for monitoring vehicle parameters. The authors of this paper have used a CAN bus protocol which is a serial field bus system which reduces wiring cost in automobile technology. The vehicle parameters chosen are speed, engine temperature, and brake status measurement. The system uses three nodes for communication using a CAN bus and accessed modern vehicle parameters from ECU using OBD-II connector. In this system, the two-node system senses the parameters using sensors. The Hall-effect sensor is used for speed measurement; a simple switch is placed below the brake pedal for brake detection, and these two measured parameters are passed through the third node using the CAN bus. The third node can be used to access all the parameters from the ECU, and the system can be used to control the vehicle remotely. The MCU used is Raspberry Pi as a single-board minicomputer and a SPI connector for transmitting CAN frames and for communication with a CAN controller on Raspberry Pi. Thus, developing a three-node network system for accessing the vehicle information like the speed, break status etc., from the ECU. It is crucial to understand the effects on distraction and usability for driving using an analog dial and a digital one.

This paper [4] comprises a study based on an analog versus digital speedometer in simulated truck driving. The authors of this paper have speedometer as a parameter since its design has severe implications on road safety. The analog speedometer was composed of a semi-circular fixed scale of 10 km/h ranging from 0 to 120 km/h, and the digital speedometer was a two digit with a refresh rate of 1.25 Hz. Three tasks were performed for conducting the experiment. In task one, the drivers were informed to report the speed. In task two, the participants were asked to inform if the speed value was above or below 50 km/h, and in the third task, the drivers were asked to say if the speed was decreasing or increasing; the experiment was conducted for six sessions corresponding to three tasks, and two speedometers display one being analog and the other digital, respectively. Each participant, after the end of the study, was asked the preference and was individually tested for all the conditions. The scores were

evaluated using the non-parametric Friedman and Wilcoxon tests for paired samples along with the visual distraction which analyzed the total glance time away from the road. The results show that for task one and two; the digital speedometer showed clear advantages. The readings showed that the reported cognitive effort was much lower for the digital speedometer as compared to the analog speedometer. On the contrary, for task three, the analog speedometer was more appropriate. Regarding the usability, both the speedometers were considered good, but the propensity to error was perceived differently; the drivers reported that they feel they make more errors using the analog speedometer. Hence, results of the comparative study between the analog and digital speedometers for truck driving show that the drivers favored digital speedometers more than analog.

3 Methodology

The implementation of this project is in three parts: model-in-loop, software-in-loop, and hardware-in-loop, respectively. Figure 1 shows the block diagram. The inputs required are fetched from the OBD-II adapter of the car. An OBD-II is then connected to the microcontroller that is Arduino UNO using a Bluetooth module HC-05. Once the connection is established, the real-time diagnostics are fetched and displayed on the LCD screen. For displaying, the values MATLAB's App Designer is used for designing a graphical user interface for vehicle dashboard.

Figure 2 shows the MIL, SIL, HIL methodology of our project. Model-in-loop or MIL is the initial stage and is run in the native development environment which

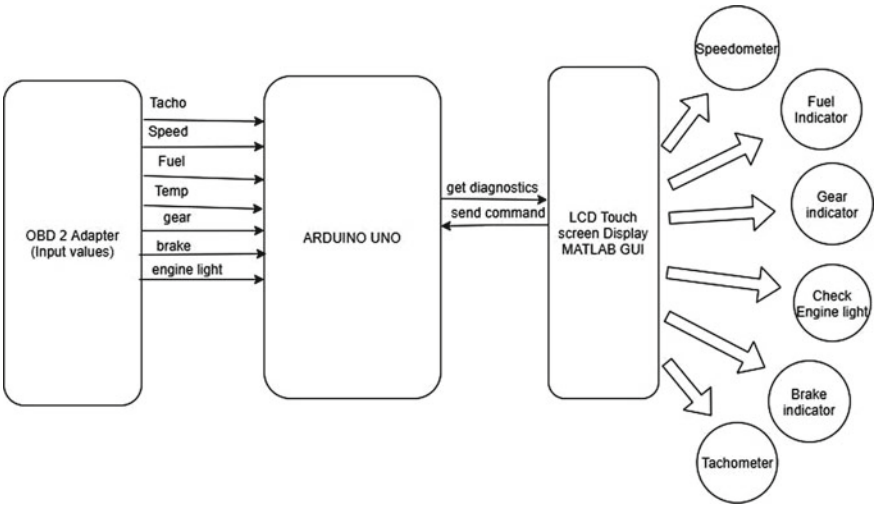


Fig. 1 Block diagram

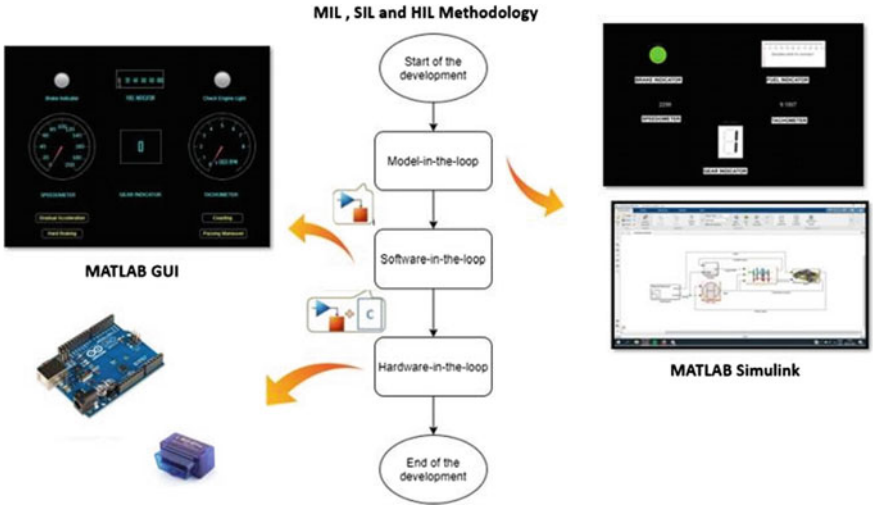


Fig. 2 MIL, SIL, HIL methodology

is Simulink for our project. The connection between subsystem components takes the form of a signal bus regardless how the signals can be physically implemented. Simulations are run on Simulink to obtain visual results on a virtual dashboard which is one of the provided features of Simulink. Inputs given to the simulations are brakes and throttle. We have provided four scenarios, namely passing maneuver, gradual acceleration, hard braking, and coasting, to test different types of braking and throttle and obtain values for those situations. All the results are obtained and can be displayed on analog gauges or digital displays. We are using digital displays for displaying real-time information. SIL stage or software-in-loop stage is the next stage after MIL. Components of the model remain the same, but the only difference is, the model is compiled into C language or Python and run with a fixed time step instead of a processor time which was used for simulations during MIL stage. This is done by designing GUI which we have made using MATLAB App Designer. Values obtained from the mentioned simulations are run on this GUI and displayed. Design is taken care of in such a way it is easily readable by the driver and doesn't hurt their eyes. After successful simulations, in the next stage, i.e., HIL, we assemble our prototype using processors and connect it to the OBD-II connector of the car. The LCD is connected to the dashboard of the car, and all the results are displayed on it. The GUI is uploaded onto the LCD and is interfaced with a microcontroller which is used to take values from the OBD port. OBD-II values can be accessed using PIDs or parameter IDs which are unique to each parameter but maintained as an international standard. These values will be displayed real time. The parameters are speedometer, tachometer, brake indicator, gear indicator, and ambient temperature.

3.1 Model-in-Loop

The vehicle model is first simulated on a simulation platform using Simulink. For this, we referred to the vehicle car dashboard model available with the MathWorks Simulink. The model consists of circular gauges for the dashboard and different user inputs to model different vehicle maneuvers. For MIL, there are two parts to the circuit. Speedometer, brake indicator, gear indicator, and tachometer have a separate circuit, and ambient temperature has a separate circuit. Figure 3 represents an automatic transmission four-gear car which has inputs as brake and throttle in the Signal Builder block.

This circuit has been given four different types of scenarios with different brake and throttle inputs as mentioned before. The brake and the throttle inputs are gradual acceleration; as can be seen from Fig. 4, no brakes are applied; the car is gradually accelerating. The passing maneuver is when a car tries to overtake another vehicle in front of them in a lane. It entails a high risk of collision with vehicles traveling in the opposite direction. Hard braking is another input signal given to the model which indicates sudden pressing of brakes usually in situations where you need to stop the vehicle immediately to avoid any mishap. Coasting while driving is defined as traveling either with the clutch pedal held down or with the car in neutral. Figure 5 is the circuit for displaying ambient temperature of the car. By default, it displays the temperature outside the car. After taking user input, it adjusts the temperature of the car near the user input value.

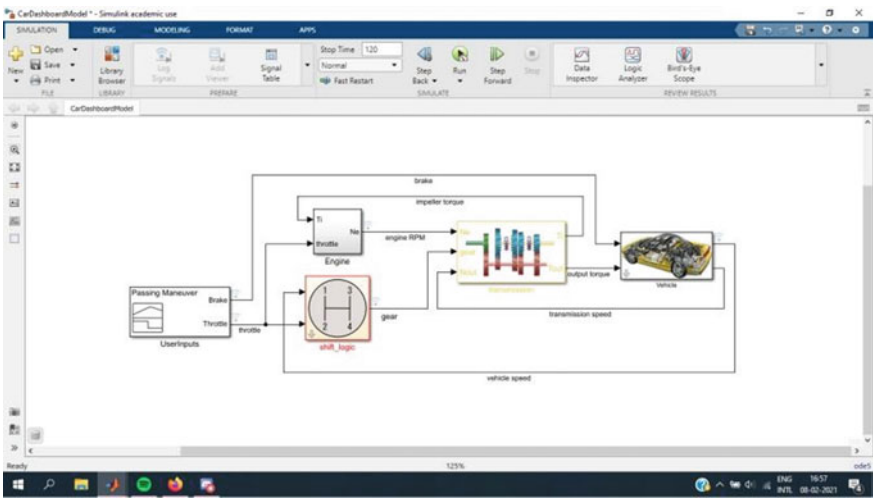


Fig. 3 Circuit for speedometer, brake indicator, and tachometer

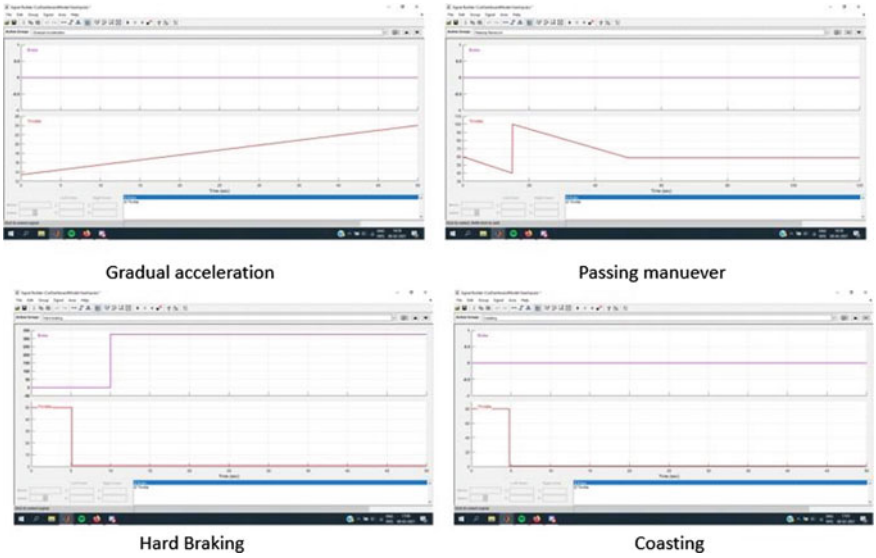


Fig. 4 Brake and throttle inputs where top graph is for brake inputs and bottom graph is for throttle inputs, both varying with time

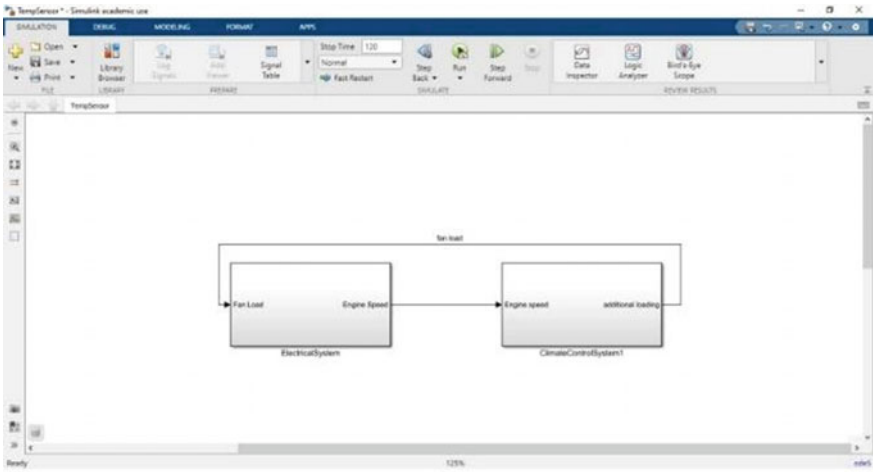


Fig. 5 Ambient temperature model

3.2 Software-in-Loop

For SIL, we designed a GUI on MATLAB's App Designer using the MATLAB version R2020b [5]. Graphical user interface (GUI) is an interface that helps a user to communicate with the machine. The newest method for creating a graphical user

interface is programmable using the graphical user interface developer environment (GUIDE) and the App Designer environment which was published in 2016. We have used the App Designer which is the newest way in MATLAB to build applications and is released in the version R2016a. There are two windows for building an application, namely the layout window and the code view. For making the layout, the visual components like gauges, lamps, buttons, switches, and knobs can be added on the window using drag and drop method. In the code view, the functionalities of the program can be handled by programming an object-oriented code. Values from above simulations are obtained from the MIL model by using Simulink's data inspector and is exported onto an Excel sheet and given as inputs to the GUI for displaying. The dashboard is created taking into consideration the proper gauges required for particular parameters to display and encoded with the values obtained from the abovementioned Simulink simulations. Keeping the aesthetics in mind, color scheme and visuals were selected. The layout and the functionalities of the entire application were programmed using MATLAB functions in the code editor. A C code was written in MATLAB's App Designer to fetch values from the dataset and to display it on the dashboard created. A callback function which is a function pointer to another function is used to communicate with the user interface control to style, position, and to set values. The dashboard consists of two circular gauges to display speedometer and tachometer values, a linear gauge for displaying the fuel indicator values, a digital display for gear indicator, and two lamps for checking engine light and brake indicator, respectively. There are four buttons for all four scenarios. The values in the dataset are given for all the four cases for the chosen vehicle parameters, respectively. Figure 6 shows the dashboard created using MATLAB's App Designer.

Fig. 6 Vehicle dashboard created using MATLAB's App Designer

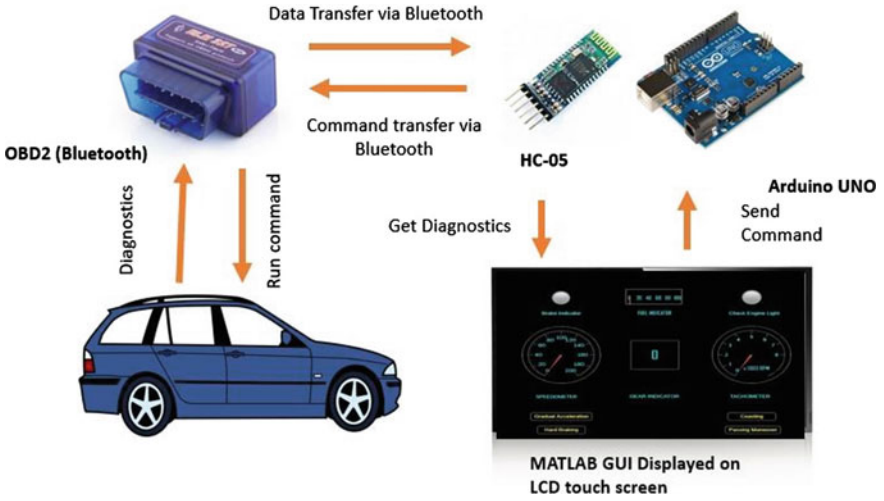


Fig. 7 Working of the model

3.3 Hardware-in-Loop

For hardware-in-loop [6], hardware components used are OBD-II ELM327 Bluetooth adapter, HC-05 Bluetooth module, Arduino UNO as microcontroller, and LCD touch screen display. These will interact with the vehicle's dashboard [7]. Figure 7 shows the working of the model. The ELM327 is used for translating the onboard diagnostics (OBD) interface found in most modern cars. The ELM327 is a PIC microcontroller produced by ELM electronics. ELM327 supports the standard OBD-II protocols: SAE J1850 PWM, SAE J1850 VPW, ISO 9141-2, ISO14230-4, and ISO 15765-4. It is available in two voltage ranges ELM327 v2.3 and ELM327L v2.2. The functions of an ELM327 is to read diagnostic trouble codes, both generic and manufacturer-specific, to clear some trouble codes and turn off the malfunction indicator light or check engine light, and to display current sensor data. The following values can be obtained using ELM327: engine RPM, calculated load value, coolant temperature, fuel system status, vehicle speed, short term, fuel trim, long-term fuel trim, intake manifold pressure, timing advance, intake air temperature, air flow rate, absolute throttle position, oxygen sensor voltages/associated short-term fuel trims, fuel system status, and fuel pressure. The Arduino UNO is used as a microcontroller board which is based on the microchip ATMega328P microcontroller and developed by Arduino. HC-05 Bluetooth module has full-duplex wireless functionalities which can communicate with microcontrollers like Arduino UNO or any device with Bluetooth functionality. The HC-05 has two operating modes: One is the data mode where it can send and receive data from other Bluetooth devices, and the other is the AT command mode which can be used to change the default settings. It is used for establishing a connection between Arduino UNO with serial UART and is configured as a master

to connect to ELM327 OBD-II Bluetooth which is the slave. HC-05 is connected to a 5 V power and ground and to the TX/RX pins of the Arduino. The initialization with AT commands begins as soon as the connection is established with ELM327. After that we can successfully read the values from the car's ECU.

4 Results

A prototype of the vehicle cluster panel can be made by following the MIL, HIL, and SIL methodology. In the MIL part, the simulation of the vehicle cluster was achieved by giving four different signals for testing the dashboard on Simulink. In SIL, we successfully designed and coded a graphical user interface for the dashboard by using MATLAB's App Designer. For HIL, we figured out the way in which the hardware components can be integrated to make a successful dashboard for a vehicle. This system eliminates the parallax errors caused by analog dashboards, and accurate readings can be obtained. Furthermore, the graphical user interface has been designed using MATLAB's App Designer which needs to be compatible with the LCD screen and is designed taking into consideration the appropriate user interface for the driver to have less distractions.

5 Conclusion

Hence, following the methodology for this research has benefited us to efficiently complete the project pertaining to this research. The methodology we followed helped in a sequential approach to tackle the problems faced at every stage to make our project more efficacious. The categorization of the model into smaller modules aided us to simplify it, but the transition from one phase to the other was challenging. We faced hardware failures, especially in the Bluetooth module which resulted in in choosing HC-05 version 2 which wasn't readily available. The difficulties faced were majorly because of the version of the Bluetooth module. It was figured that the AT commands for the latest version of the HC-05 Bluetooth module faced issues. This problem can be overcome by using the older version of HC-05 or HM-10 Bluetooth modules. Hence, this research work can be implemented to make a durable digital dashboard for automobiles since the advantages of having a digital dashboard are more than analog dashboards, and it could be accessible to the customers at lower costs.

References

1. Nibert J, Herniter ME, Chambers Z (2012) Model-based system design for MIL, SIL, and HIL. World Electr Veh J 5(4):1121–1130. https://doi.org/10.3390/wevj5041121
2. Pawlenka T, Kulhanek J, Tomčík P, Zapletal R (2019) Design of digital CAN based car dashboard unit. In: 2019 20th international Carpathian control conference (ICCC), pp 1–4. https://doi.org/10.1109/CarpathianCC.2019.8766024
3. Salunkhe AA, Kamble PP, Jadhav R (2016) Design and implementation of CAN bus protocol for monitoring vehicle parameters. In: 2016 IEEE international conference on recent trends in electronics, information communication technology (RTEICT), pp 301–304. https://doi.org/10.1109/RTEICT.2016.7807831
4. Francois M, Osiurak F, Fort A, Crave P, Navarro J (2016) Analogue versus digital speedometer: effects on distraction and usability for truck driving. In: European conference on human centred design for intelligent transport systems. Loughborough, France, 8 p
5. Grepl R (2011) Real-time control prototyping in MATLAB/Simulink: review of tools for research and education in mechatronics. IEEE Int Conf Mechatronics 2011:881–886. https://doi.org/10.1109/ICMECH.2011.5971238
6. Huang Y, Mouzakitis A, McMurran R, Dhadyalla G, Jones RP (2008) Design validation testing of vehicle instrument cluster using machine vision and hardware-in-the-loop. In: 2008 IEEE international conference on vehicular electronics and safety, pp 265–270. https://doi.org/10.1109/ICVES.2008.4640857
7. Hegde SB, Mohan Kumar NK, Kumar M, Kumari M, Jha RK (2017) Automated driver-vehicle control system using CAN protocol. J Manage Eng Inf Technol (JMEIT). https://doi.org/10.5281/zenodo.821238

Technology for Creation of False Information Fields for UAV Protection

Elena Basan and Nikita Proshkin

Abstract The issues of ensuring the safety of UAVs are quite acute in modern times. UAVs can be used to implement critical processes. Such processes can be relaying communications, searching for missing persons, reconnaissance operations, guarding facilities, and much more. At the same time, the UAV is often outside the controlled area and is physically unprotected. It is difficult to protect UAVs and communication channels between UAVs since the wireless communication channels used for UAVs are physically unprotected. There are many options for physical protection that involve imposing additional noise on a channel. In this case, an attacker can try to carry out an information attack or, for a start, simply detect the UAV. In our research, we propose a way to hide UAVs by creating false information fields. Next, we test our method by analyzing the radio spectrum and comparing our fake fields with legitimate ones. The results showed the effectiveness of the developed software, which allows you to create fake access points that can be detected by an intruder and allow you to hide the real transmission.

Keywords UAV · Radio signal · Wireless communication · Defense methods · Attacks · Vulnerabilities · Software

1 Introduction

UAVs today are becoming an increasingly popular solution for performing various tasks [1]. Moreover, such tasks are often critical, for example, searching for missing people, extinguishing fires, border control, etc. [2]. At the same time, UAVs are quite vulnerable to attacks by an intruder since they are physically unprotected [3]. The most frequently implemented attacks exploit communication channel vulnerabilities. You can try to hide the UAV by providing some physical protection [4].

Using communication protocols, one can achieve that via a single channel, it is possible to transmit various types of information described above with respect

E. Basan (✉) · N. Proshkin
Southern Federal University, Chekov St., 2, 347922 Taganrog, Russian Federation
e-mail: ebasan@sfedu.ru

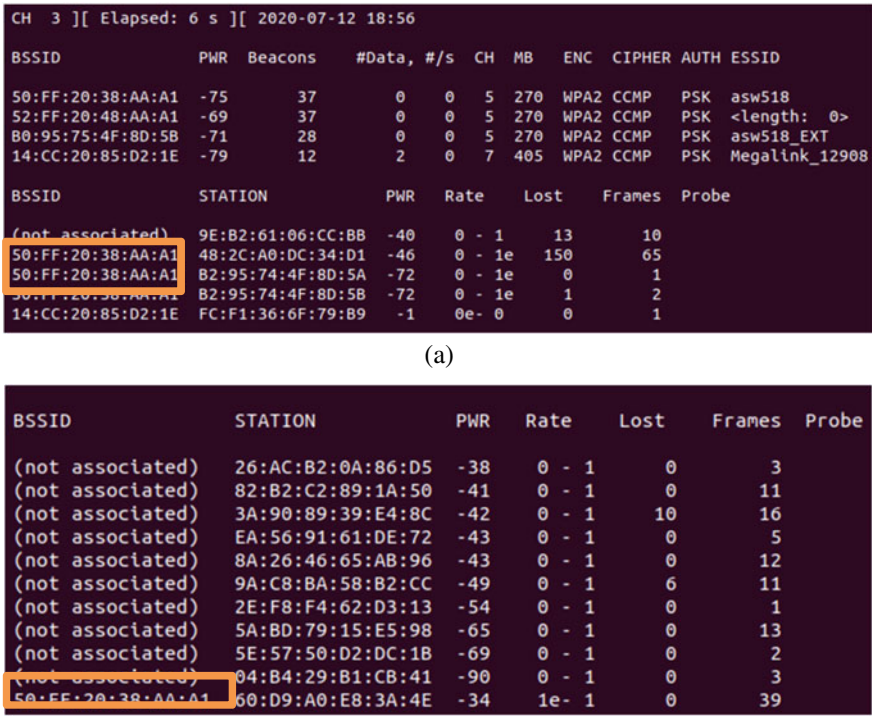
A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_72

to encoding and decoding of radio software. Considering the necessary characteristics of the UAV operation, the scenario of the simulator operation should be as follows: For example, a transmitter/receiver with a range of 150 km [C-130] uses eight channels to provide communication; therefore, with a small UAV configuration (telemetry and commands), it is possible to simulate up to four devices operating simultaneously. This method of imitation at the level of communication channels allows you to create a closeness of objects due to listening at one frequency and finding many electromagnetic waves on it. When choosing a transmitter/receiver with many communication channels, you can simulate a larger number of objects. In some models of receivers/transmitters, one channel is used, but in full duplex mode, this allows you to allocate not two channels to one device but only one, which doubles the simulation of devices. In wireless networks, Wi-Fi networks carry out similar attacks but for different purposes. For example, a beacon-flood attack simulates the presence of many neighboring access points, which should make it difficult for a client to reach a legitimate access point. Moreover, this attack does not actually affect legitimate users. A legitimate access point has a poppy address of 50: FF: 20: 38: AA: A1. Let's carry out a beacon-flood attack and try to find our legitimate access point. Figure 1 shows that in addition to a legitimate access point, there are many others without an identifier in this case.

2 The Software for Creating a False Image of a UAV

Often, the radio silence mode is used to hide the UAV. This mode does not provide for the possibility of transmitting information to the UAV. At the same time, considering the current trends in the creation of a united cyberspace, when devices must not only collect and transmit data but also communicate with each other to coordinate their actions, working in radio silence mode may be unacceptable [5]. Methods of imposing radio noise on the communication channel can also be used, which masks the legitimate transmission of information [6]. This method requires additional equipment and costs, so it may not always be applicable [7].

The proposed software module provides for the concealment of a legitimate communication channel by forging a variety of wireless access points. The study made it possible to determine the parameters of civil UAVs for network communication and to emulate these parameters to hide a legitimate UAV [8]. Civilian UAVs work as follows: the UAV's network adapter is switched to the mode of creating an access point, after the operator has created an access point, knowing the connection parameters (as a rule, this is the MAC address of the access point and its identifier) can connect to it from an authorized device [9]. Thus, the task of the software module is to create several access points that would advise a civilian UAV in terms of parameters. When developing the software module, it was possible to change the number of created access points. This approach will minimize the risks associated with the possibility of implementing attacks over wireless communication channels of the UAV [10]. Each access point operates on its own channel; the IEEE 802.11



```
CH  3 ][ Elapsed: 6 s ][ 2020-07-12 18:56

BSSID              PWR  Beacons    #Data, #/s  CH  MB   ENC  CIPHER AUTH ESSID

50:FF:20:38:AA:A1  -75       37        0    0   5  270  WPA2 CCMP   PSK  asw518
52:FF:20:48:AA:A1  -69       37        0    0   5  270  WPA2 CCMP   PSK  <length:  0>
B0:95:75:4F:8D:5B  -71       28        0    0   5  270  WPA2 CCMP   PSK  asw518_EXT
14:CC:20:85:D2:1E  -79       12        2    0   7  405  WPA2 CCMP   PSK  Megalink_12908

BSSID              STATION            PWR   Rate    Lost    Frames  Probe

(not associated)   9E:B2:61:06:CC:BB  -40    0 - 1     13        10
50:FF:20:38:AA:A1  48:2C:A0:DC:34:D1  -46    0 - 1e   150        65
50:FF:20:38:AA:A1  B2:95:74:4F:8D:5A  -72    0 - 1e     0         1
                   B2:95:74:4F:8D:5B  -72    0 - 1e     1         2
14:CC:20:85:D2:1E  FC:F1:36:6F:79:B9   -1    0e- 0      0         1
```

(a)

```
BSSID              STATION            PWR   Rate    Lost    Frames  Probe

(not associated)   26:AC:B2:0A:86:D5  -38    0 - 1      0         3
(not associated)   82:B2:C2:89:1A:50  -41    0 - 1      0        11
(not associated)   3A:90:89:39:E4:8C  -42    0 - 1     10        16
(not associated)   EA:56:91:61:DE:72  -43    0 - 1      0         5
(not associated)   8A:26:46:65:AB:96  -43    0 - 1      0        12
(not associated)   9A:C8:BA:58:B2:CC  -49    0 - 1      6        11
(not associated)   2E:F8:F4:62:D3:13  -54    0 - 1      0         1
(not associated)   5A:BD:79:15:E5:98  -65    0 - 1      0        13
(not associated)   5E:57:50:D2:DC:1B  -69    0 - 1      0         2
                   04:B4:29:B1:CB:41  -90    0 - 1      0         3
50:FF:20:38:AA:A1  60:D9:A0:E8:3A:4E  -34    1e- 1      0        39
```

(b)

Fig. 1 Analysis of network activity. **a** Under normal conditions, **b** Under a beacon-flood attack

standard is implemented in such a way that broadcasting on a specific channel occurs all the time; therefore, the analysis of the Wi-Fi frequency with the module turned on will show the radioactivity of several devices (UAVs) [11]. The software module is implemented for a single-board computer and requires the use of an external Wi-Fi adapter to be able to create false wireless access points. The emulation module consists of a Raspberry Pi 3 model B microcomputer, a microcomputer battery, a Linux "Raspbian" OS for the Raspberry Pi, a ZyXEL G-202 EE USB wireless Wi-Fi adapter. and software that runs a script to create many fake Wi-Fi hotspots automatically when a certain event occurs [12]. The "airmon-ng" [13] utility in Linux OS is used to create fake access points [14]. The software consists of three modules: (1) The interface initialization module is required to correctly determine the network interface on which fake access points will be created. (2) The module for generating access points, serves to transfer the desired interface to the monitor mode and create access points on it. (3) The event response module acts as a connecting shell between the previous two modules, performs the functions of receiving and transmitting data from one module to another and calling them, and turns on automatically under a certain condition and two text files: (1) configuration file of the initialization module; (2) the log file of the generation module.

3 Analysis of the Frequency Spectrum Emitted by the Module for Creating False Information Fields

The experiments were carried out using a spectrum analyzer GW Instek (GSP827) under conditions of low influence of powerful radiating antennas. The laboratory stand is shown in Fig. 2 [15].

2.4 GHz wireless standard allows only 14 channels with 20–22 MHz channel width. Channels 1, 6, 11 are optimal for simultaneous use; 2, 7, 12; 3, 8, 13 or 4, 9, 14. But, in this study, it is believed that there is no need to transfer useful data over them. Active radio traffic is generated by additional emulation flags. Therefore, all channels can be used. The standard for 5 GHz has 140 frequency-separated channels, respectively; 10 times more APs can be deployed, but the radiation radius is halved. Therefore, we carried out an experiment with a frequency of 2.4 GHz, with a radiation radius of up to 150 m in an open area with a transmitter power of 18 dBm. The maximum number of 14 channels at 2.4 GHz does not mean that only 14 APs can be deployed. Two or more APs operating on the same channel are simply superimposed on each other and emit traffic in turn. In the legitimate implementation of many APs, of course, this will lead to a loss of communication efficiency, but we do not care. In the radio frequency method of emulating several UAVs, on the contrary, active radio traffic on the radio radar is encouraged, which is like an active TCP connection and packet transmission.

3.1 Frequency Spectrum Analysis of 2.4 GHz Real AP

First, let's analyze the frequency spectrum of a working phone as an AP, but without connecting any devices to it that is the AP sends Beacon packets with a certain period, but there is no radio traffic. As you can see on the analyzer display, the maximum signal level reaches approximately −30 dBm with a frequency of approximately 2.46 GHz, which corresponds to channel 11 in the Wi-Fi standard. The frequency of occurrence of carriers is about two times per second. Figure 4 shows the connection of another device to the phone's AP; a TCP connection has been established between the

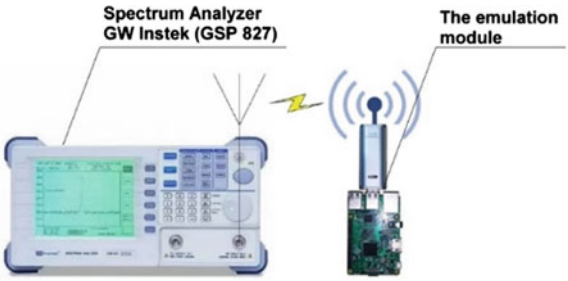


Fig. 2 Experimental stand for spectrum analysis

device and the Web server, and packets are being actively transmitted. The difference between sending beacon packets and an active TCP connection on the spectrum analyzer is that the signal width and the frequency of occurrence of carriers have increased; the signal has become much continuous; this indicates that some device is communicating via the wireless network at a particular moment with another device. Now that we understand what the signal of a simple AP and AP looks like with a device connected to it and an active exchange of packets, let's move on to analyzing fake APs; how the quantity and separation by channels affects the type of signal, the frequency of occurrence of peaks, and the width of radio traffic in general.

3.2 Frequency Spectrum Analysis of Fake APs

In our experiment, we recreated 25 APs, increasing their number gradually, without turning off the previous ones. The signal is completely like the signal in Fig. 4, but with a lower level equal to −50 dBm due to the lower transmitter power, channel 12 was set, which corresponds to the average frequency of the signal. Figure 3 shows the spectrum of the one AP signal, but with the active emulation, flags turned on. The frequency of the peaks increased by about two times, and the signal level slightly increased by 0.2–0.4 dBm. Another carrier has appeared, which was created by the option of sending additional probe packets for known devices. The spectrum bandwidth is specified at 75 MHz.

Figure 4 shows the signal spectra of different numbers of fake access points. It is clearly seen that the number of carriers increases; the distance between them decreases, and traffic becomes more active due to the increase in the number of access points.

Figure 5 shows the dependences of the parameters: signal level, number of carrier,, and frequency of occurrence of carriers on the number of access points. The ZyXEL Wi-Fi adapter can only broadcast 14 channels each 20–22 MHz wide. We use our software to set channel numbers 1, 6, 11, 16, 21, 26, 31, 36, 41, 46, etc. Therefore, a channel with a number greater than 14 will have a number calculated by the Formula 1:

$$Nk = NPk \mod 14 \tag{1}$$

where *Nk* is the actual channel number, *NPk*—programmable channel number, mod is the integer remainder of division.

From this, it follows that our actual channels overlap each other. However, since the adapter cannot broadcast simultaneously on 14 channels, broadcasting occurs alternately, but the queue does not consist of actual channels, but of the channels that we specified programmatically. That is, the transmission of packets occurs first on 1, then on 6, then on 11, on 2, on 7 channels, etc. Due to this organization of the queue of sending packets, the interference of the waves is much less. Figure 6 shows

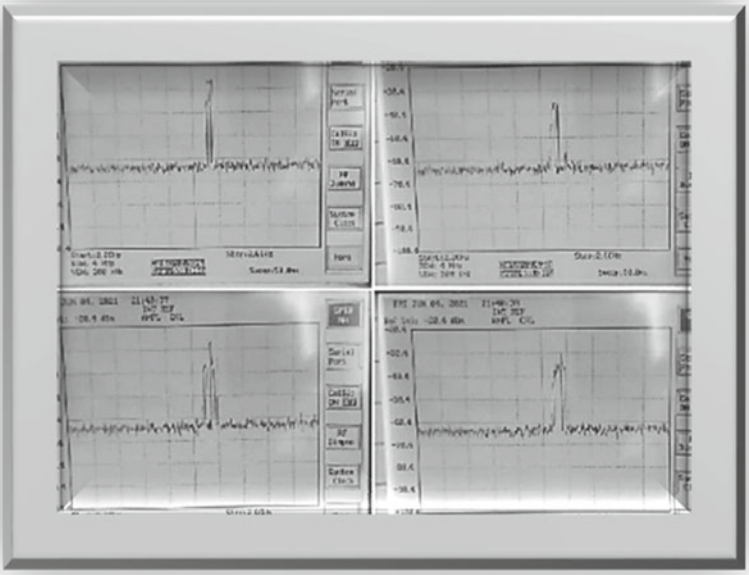


Fig. 3 Spectrum of the signal under TCP connections

a comparison of a real-TCP connection at 2.4 GHz and five fake APs. The form of the signal parameters is like each other, which will give an erroneous understanding of the reconnaissance object.

4 Conclusion

In this study, we have demonstrated the operation of our radio frequency emulation module for several UAVs by creating several fake APs that transmit traffic consisting only of beacon and probe packets. Analysis of the spectra of different alternately increasing number of APs showed that with an increase in the number of APs, the number of carrier frequencies changes, which shows several working separate Wi-Fi devices, the frequency of occurrence of peaks increases—the likelihood of enemy confusion increases—emulation of active radio traffic, and what is more 9 simultaneously working APs do not change the efficiency of entanglement in any way. Analysis of the spectrum of the real operation of the device showed that when 13 APs are simultaneously turned on, the signal becomes more continuous and effectively entangled compared to the 9th AP.

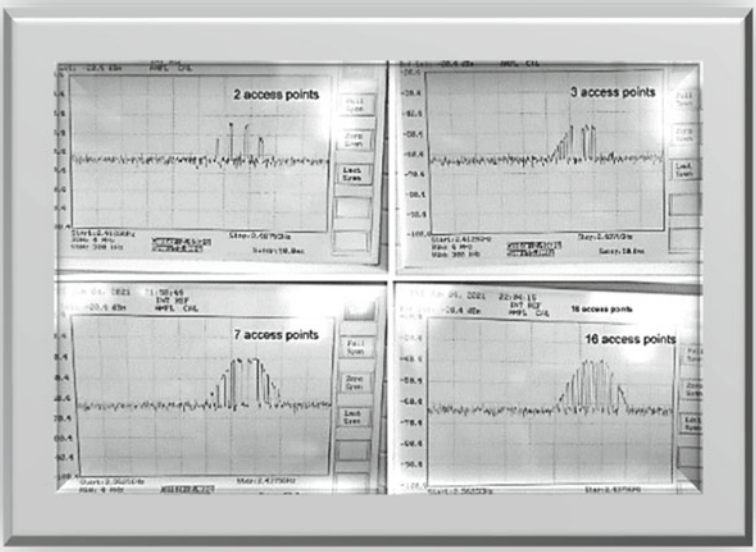


Fig. 4 Signals of a different number of APs

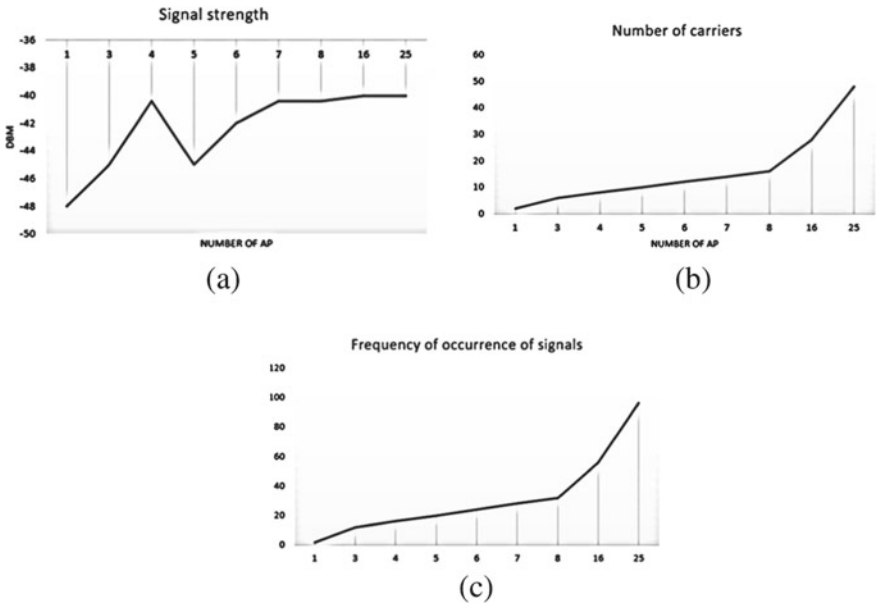


Fig. 5 Dependence of **a** the signal level on the number of APs **b** the number of carriers on the number of APs **c** the frequency of occurrence of carriers on the number of APs

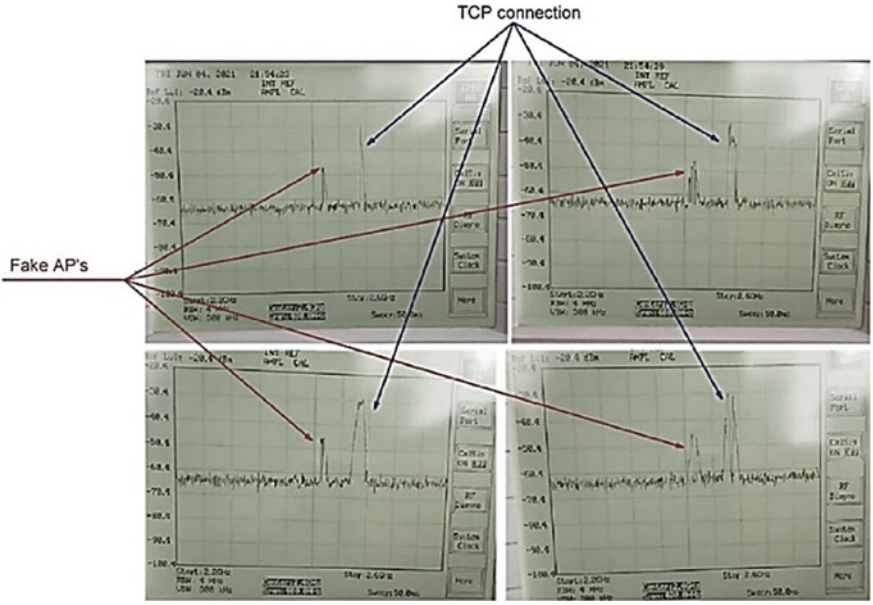


Fig. 6 Comparison of the signals of five APs and a TCP connection

This is since with the alternate increase in the number of APs, we make an error in the time of launching the next AP; thereby, we can indicate the launch of the AP at the wrong time. When 13 APs are created at the same time and launched, each AP broadcasts at an equal time interval from the neighboring one; thus, the effect of continuous radio traffic is obtained. In conclusion, I would like to note that we conducted an experiment with the simplest Wi-Fi adapter and a weak antenna. In real-life conditions, a much more powerful emitter should be used to increase the range and signal level, and a more sensitive antenna will be used by the enemy. Also, to increase the range and signal level, you can use narrow-beam antennas; however, it is necessary to accurately understand the location of the enemy's reconnaissance.

Acknowledgements This research was funded by the Russian Science Foundation grant number 21-79-00194, https://rscf.ru/project/21-79-00194/., in Southern Federal University.

References

1. Xu C, Liao X, Tan J, Ye H, Lu H (2020) Recent research progress of unmanned aerial vehicle regulation policies and technologies in urban low altitude. IEEE Access 8:74175–74194. https://doi.org/10.1109/ACCESS.2020.2987622
2. Shi Y, Bai M, Li Y (2018) Study on UAV remote sensing technology in irrigation district informationization construction and application. In: 10th international conference on measuring

technology and mechatronics automation (ICMTMA). Changsha, China, pp 252–255. https://doi.org/10.1109/ICMTMA.2018.00067
3. Gao X, Jia H, Chen Z, Yuan G, Yang S (2020) UAV security situation awareness method based on semantic analysis. In: 2020 IEEE international conference on power, intelligent computing and systems (ICPICS). Shenyang, China, pp 272–276. https://doi.org/10.1109/ICPICS50287.2020.9201954
4. Hooper M et al (2016) Securing commercial Wi-Fi-based UAVs from common security attacks. In: MILCOM 2016—2016 IEEE military communications conference. Baltimore, MD, USA, pp 1213–1218. https://doi.org/10.1109/MILCOM.2016.7795496
5. Mikhalevich IF, Trapeznikov VA (2019) Critical infrastructure security: alignment of views. In: Systems of signals generating and processing in the field of on board communications. Moscow Technical University of Communications and Informatics, Russia, pp 1–5. https://doi.org/10.1109/SOSG.2019.8706821
6. Ilioudis CV, Clemente C, Soraghan J (2019) Understanding the potential of self-protection jamming on board of miniature UAVs. In: International radar conference (RADAR). Toulon, France, pp 1–6. https://doi.org/10.1109/RADAR41533.2019.171405
7. Li X, Ju R, Wang H, Sun Y (2018) The design and research of data transmission on remote radio control in different noise channel. In: 13th world congress on intelligent control and automation (WCICA). Changsha, China, pp 1306–1311. https://doi.org/10.1109/WCICA.2018.8630340
8. Proshkin N, Basan E, Lapina M (2021) Radio frequency method for emulating multiple UAVs. In: 2021 17th international conference on intelligent environments (IE). Dubai, United Arab Republic, pp 1–4. https://doi.org/10.1109/IE51775.2021.9486599
9. Deligne E (2012) ARDrone corruption. J Comput Virol 8:15–27
10. Astaburuaga I, Lombardi A, La Torre B, Hughes C, Sengupta S (2019) Vulnerability analysis of AR.Drone 2.0, an embedded Linux system. In: IEEE 9th annual computing and communication workshop and conference (CCWC). United States, pp 0666–0672. https://doi.org/10.1109/CCWC.2019.8666464
11. Caballé MC, Augé AC, Lopez-Aguilera E, Garcia-Villegas E, Demirkol I, Aspas JP (2019) An alternative to IEEE 802.11ba: wake-up radio with legacy IEEE 802.11 transmitters. IEEE Access 7:48068–48086. https://doi.org/10.1109/ACCESS.2019.2909847
12. Madruga S, Tavares A, Brito A, Nascimento T (2018) A project of an embedded control system for autonomous quadrotor UAVs. In: Latin American robotic symposium, 2018 Brazilian symposium on robotics (SBR) and 2018 workshop on robotics in education (WRE). João Pessoa, Brazil, pp 483–489. https://doi.org/10.1109/LARS/SBR/WRE.2018.00091
13. Carranza A, Mayorga D, DeCusatis C (2018) Comparison of wireless network penetration testing tools on desktops and Raspberry Pi platforms. In: Proceedings of the 16th LACCEI international multi-conference for engineering, education and technology. Boca Raton, Florida, USA, pp 1–5
14. Abril-García JH et al (2019) Application to monitoring a USB control with Python in windows, Mac OS and Raspbian OS. ECORFAN J Democratic Repub Congo 5(8):1–6
15. Korneev S (2008) Digital spectrum analyzer GSP-7830 manufactured by GW Instek. Compon Technol 1(78):158–162

Implementation of CSRF and XSS Using ZAP for Vulnerability Testing on Web Applications

Farzana Sultana, Md. Mynul Islam, Md. Tanbir Hasan, and Md. Ismail Jabiullah

Abstract The security of Web applications is one noteworthy component that is often overlooked within the creation of Web apps. Web application security is required for securing websites and online services against distinctive security threats. The vulnerabilities of the Web applications are for the most part the outcome of a need for sanitization of input/output which is frequently utilized either to misuse source code or to pick up unauthorized access. An attacker can misuse vulnerabilities in an application's code. The security of Web applications may be a central component of any Web-based commerce. The security of Web applications deals particularly with the security encompassing websites, Web applications, and Web administrations such as APIs. This paper gives a testing approach for vulnerability evaluation of Web applications to address the extent of security issues. We illustrate the vulnerability assessment tests on Web applications. Showing how with an aggregation of tools, the vulnerability testing broadcast for Web applications can be enhanced.

Keywords Owaspbwa · ZAP · OWASP · Penetration testing security · Vulnerability assessment · Web server · XSS · CSRF

F. Sultana (✉) · Md. Mynul Islam · Md. Tanbir Hasan · Md. Ismail Jabiullah
Daffodil International University, Dhaka, Bangladesh, India
e-mail: farzana15-1624@diu.edu.bd

Md. Mynul Islam
e-mail: mynul15-10403@diu.edu.bd

Md. Tanbir Hasan
e-mail: tanbir15-10420@diu.edu.bd

Md. Ismail Jabiullah
e-mail: drismail.cse@diu.edu.bd

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_73

1 Introduction

The proceeded development of Internet usage as well as the development of Web applications frontal area the challenges of IT security, especially in terms of information privacy, data integrity, and benefit availability.

It was predicted on the ongoing state of the craftsmanship on security and all the security reports like OWASP Best
Ten 2013, CWE/SANS 25, and WhiteHat Site Security Measurement Report 2013 [1]. The applications of the Web are the primary well-known targets when talking of cyber-attacks. These days, Web servers are getting to be uncommonly vital in various applications, particularly in information technology-based organizations, mechanical embedded contraptions utilized in mechanical plants, conjointly in electrical utilities. Attackers can enter through the security escape clauses of a Web server utilized in conventional IT applications and interior of the implanted gadgets utilized in basic mechanical situations to compromise its usefulness, accessibility, productivity, and tastefulness. An entrance test endeavors to effectively misuse shortcomings in an environment. Whereas, a helplessness filter can be robotized; an entrance test requires different levels of ability. Considering such fundamentals, testing the security aspects of a website is exceptionally critical. There are physical and mechanized strategies enacting the security appraisals of the application of the Web. Within the mechanization strategy, there is a part of penetration testing instruments, which are accessible either as open-source or as a commercial item with distinctive functionalities and applicability. It is conceivable to attain thorough helplessness appraisal tests that can cover most of the vulnerabilities.

In this paper, we bargain a correlative and combined analysis of the Web application vulnerability assessment attack. It provides the related work associated with the Web application vulnerability assessments. This paper provides a scanning approach for the vulnerability assessment of a Web application. At last, it concludes with a viable approach towards vulnerability evaluation tests. Vulnerability scanners alarm companies to the preexisting defects in their code and where they are found. Penetration tests endeavor to misuse the vulnerabilities in a framework to decide whether unauthorized access or other malicious action is conceivable and recognize which imperfections posture a risk to the application. Penetration tests discover exploitable imperfections and degree the seriousness of each. A penetration test is implied to appear how damaging a flaw may well be in a genuine attack instead of discovering each blemish in a system. Together, penetration testing and helplessness assessment tools give a point-by-point picture of the flaws that exist in an application and the dangers related to those blemishes. In this paper, comparison of these two approaches is given. Infiltration testing is way better compared to vulnerability appraisal because it abuses powerlessness, whereas vulnerability appraisal is superior in terms of coverage over penetration testing.

2 Testing Approach

The term "testing approach" implies that a vulnerability assessment is a methodology that is used for identifying security loopholes within an application and infrastructure and their subsequent remediation [2]. It involves intense scanning of components either by individual security experts or network security admins in an organization. A vulnerability test approach is predicted in Fig. 1 [2].

Remediate: Remediation is essentially one portion of vulnerability administration. Vulnerability administration refers to the complete handle of overseeing your network's vulnerabilities and the dangers it faces. This incorporates finding, assessing, prioritizing, and killing dangers.

Verify: Confirm that dangers have been dispensed with through follow-up reviews.

Discover: Stock all resources over the arrangement and recognize have points of interest counting working framework and open administrations to recognize vulnerabilities. Create an arranged pattern. Distinguish security vulnerabilities on a customary mechanized plan.

Assess: A disabled appreciation may be a precise survey of security shortcomings in a data framework. It assesses in the event that the framework is helpless to any known vulnerabilities, relegates seriousness levels to those vulnerabilities, and suggests remediation or relief, on the off chance that at anything point required.

Report: Vulnerability evaluation may be a handle that distinguishes dangers, and dangers that are performed through the utilization of robotized testing devices such as vulnerability scanners. Since security vulnerabilities can permit cyber assailants to penetrate an organization's IT frameworks, it is fundamental to distinguish and thus remediate vulnerabilities, sometime recently, they can be entered and misused [2].

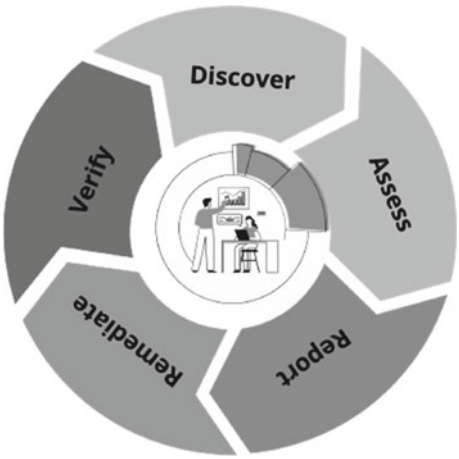


Fig. 1 Process of Vulnerability test approach

3 Scanning Process for Testing

Open-source vulnerability scanners, regularly utilized as a portion of program composition investigation (SCA) devices, are utilized to distinguish open-source components utilized in computer program ventures, and check on the off chance that security vulnerabilities are accommodated uncorrected by them, and they are proffered support reclaim organizations. The working process of scanning is given below [2]:

Scanning open-source components: The scanner audits all open-source components within the program venture, regularly by scrutinizing code stores, bundle directors, and construct devices [2].

Verifying license compliance: Most helplessness scanners can recognize computer program licenses of open-source components and confirm in the event that they struggle with organizational arrangements [2].

Identifying vulnerabilities: Scanners take information and check them against one or more databases containing information around vulnerabilities, counting common vulnerabilities and exposures (CVE) databases which are a standardized list of vulnerabilities known to security examiners and elite security explores databases [2].

4 Implementation

4.1 Testing Setup

Figure 2 shows a visual representation of the proposed testing environment of Owaspbwa (OWASP Broken Web Application Project) [1].

Owaspbwa Web server is installed on a VirtualBox. The Owaspbwa is installed in VirtualBox as an operating system that is also running an apache server and can be accessed by its IP address. Any user in that network can access it through any Web browser [1].

Figure 3 is the home page of Owaspbwa, and there are some Web applications. Every application here is for learning and testing one's skills.

4.2 CSRF

A CSRF attack is a request code where a user is strict about interacting with a page or script on a third-party site. They generate a malicious request to your site. All your

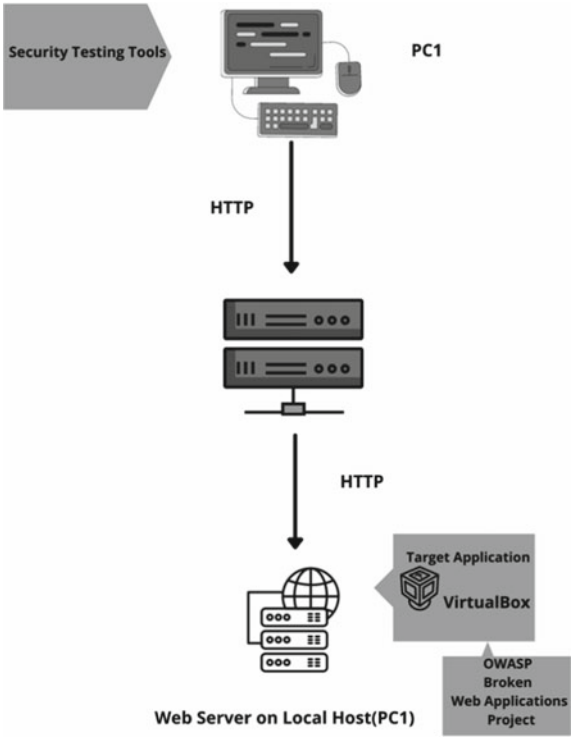


Fig. 2 The Set-up of Vulnerability test

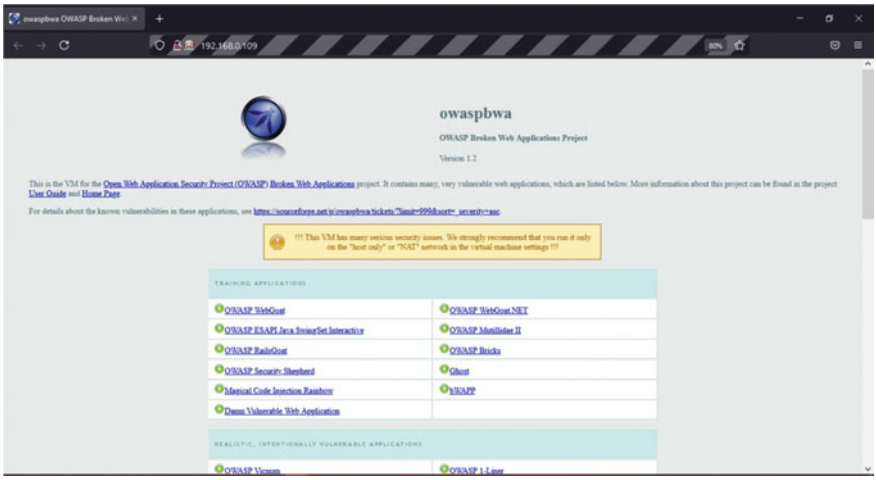


Fig. 3 Owasp broken Web application project home page

servers will see an HTTP request from an authenticated user; however, an attacker attacks control over the form of the data sent and then requests to code mischief [3].

In this Fig. 4, we change our password to see the GET request that the website uses to change the password. HTTP GET is by its exceptional nature implied to be an idempotent ask strategy. This implies that the HTTP strategy ought to not be utilized to perform state changes. Sending a GET ask ought to never cause any information to alter. Be that as it may, a few Web apps still utilize GET rather than the more suitable POST to perform state changes for operations such as changing a watchword or including a client (Fig. 5).

This Web application is employing a GET request that incorporates two parameters: the new password and confirmation of the new password. If the passwords are the same, it will carry out the request and change the password. Due to no authentication, we can just change the value of passwords and send the new link to a victim [4].

Fig. 4 DVWA cross-site request forgery page

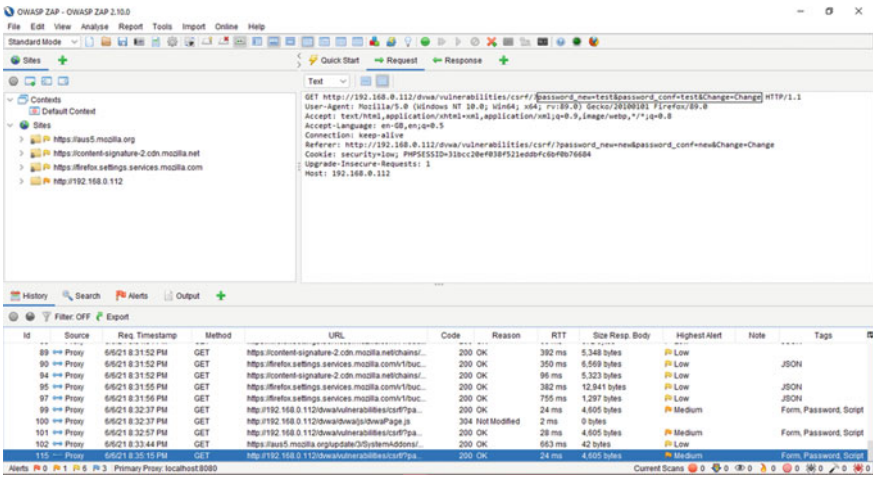


Fig. 5 Captured response on ZAP proxy

For example, if we input the following Web address of the intended Web application:

"http://192.168.0.112/dvwa/vulnerabilities/csrf/?password_new=password&password_conf=password&Change=Change", then by clicking on that link their password will change immediately.

4.3 Cross-Site Scripting (XSS)

Cross-site scripting or XSS is a consumer-side code injection attack allowing an attacker to accomplish the interactions that users have with a vulnerable application. An attacker is allowed to hype the same-origin policy by it, which is designed to lay aside several websites from each other [5]. An attacker is generally approved by cross-site scripting vulnerabilities to run a malicious script to a victim user to carry out any actions that the user is able to execute and to access any of the user's data. If the victim user has privileged access within the application, the end user's browser has no way to know that the script should not be believed; the attacker might be able to attain access to any cookies, session tokens, and other sensitive information [5].

To carry out a cross-site scripting attack, a malicious script has to be impelled into user-provided input by an attacker. Attackers can also attack by altering a request. If the Web app is vulnerable to XSS attacks, the user-provided input executes as code [5].

In Fig. 6, I will use the bWAPP from OWASP's broken Web project [1]. In this page that is shown above, it takes two inputs: first name and last name and show the welcome message as a single line output [6]. We can use that feature and enter malicious input like a simple script "< script > alert ("GG") < /script >". Also, we have to split the script because after taking the input, it makes it single line output. In the first name, we put "< script > alert ("GG")" and in the last name "< /script >". The response to the script is shown in Fig. 7.

As we can see the message inside our script showed up as a pop-up message. Now, we know it works, and we will use this to take out more information like cookies. Also, if we look at the URL http://192.168.0.108/bWAPP/xss_get.php?firstname=%3Cscript%3Ealert%28%22GG%22%29&lastname=%3C%2Fscript%

Fig. 6 Intended output

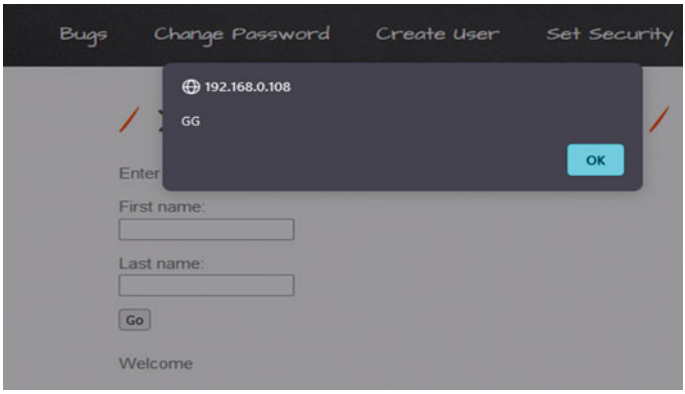


Fig. 7 XSS attack on bWAPP

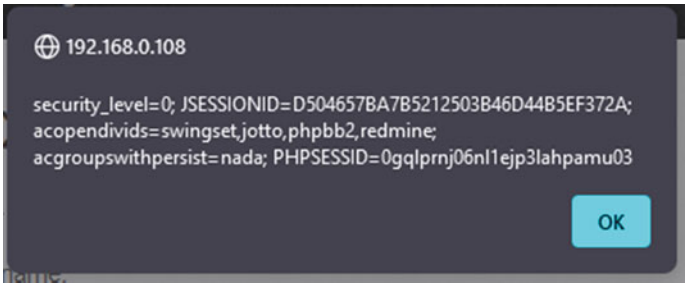


Fig. 8 Identifying security misconfiguration

3E&form=submit, we can see our script on the URL. So someone can make a victim click this link and make the script run on their browser. We can use other input like "< script > alert (document.cookie) < /script >" that will show us the cookies (Fig. 8).

An attacker can use this differently for example "< script > window.location = http://example.com:[port]cookie=+document.cookie </script >". By clicking on a link with this script, the victims will be redirected to example.com, and their cookie information will be sent to example.com. And for this example, I will use a Python server in place of example.com (Fig. 9).

And this is how an XSS attack is performed. There are also other types of XSS like stored and dom. This vulnerability is fairly common nowadays and also numerous ways to prevent it.

Fig. 9 Capturing cookies on our Python server

5 Conclusion

A process of vulnerability testing of Web applications has been designed, developed, and implemented using ZAP software that gives a trial gateway for vulnerability evaluation of Web applications to address the extent of security issues. One can illustrate the vulnerability assessment tests on Web applications. Showing how with an aggregation of tools, the vulnerability testing broadcast for Web applications can be enhanced. This can be applied in any Web-based software testing for vulnerability and penetration testing analysis.

References

1. OWASP Broken Web Applications Projects. Available from: https://sourceforge.net/projects/owaspbwa/
2. Vibhandik R, Bose AK (2015) Vulnerability assessment of web applications—a testing approach. In: 2015 forth international conference on e-technologies and networks for development (ICeND). IEEE, Lodz
3. Acunetix (n.d.) Retrieved from www.acunetix.com/websitesecurity/csrf-attacks/
4. Zeller W, Felten EW (2008) Cross-site request forgeries: exploitation and prevention. Berkeley EECS
5. ALECU F (2012) Cross site scripting (XSS) in Action. Oeconomics Knowl
6. Port Swigger (n.d.) Retrieved from portswigger.net: https://portswigger.net/web-security/cross-site-scripting

Smart Classroom: A Step Toward Digitization

Shivani Agrawal, Abhinandan Chandhok, S. Maheswari, and P. Sasikumar

Abstract The word SMART has already been implemented on a large scale in different fields from smart home to smart industry. Smart classroom service, one of the prominent technologies in the Internet of Things era, will change classrooms into being more intelligent, interconnected, and remotely controllable. The existing manual attendance system kills a tremendous amount of time for both the faculties and the students. Proxy attendance is also a huge challenge. There is a need for digitization of course materials for allowing access to it anytime from anywhere. Considerable amounts of energy can also be saved by switching off the smart board when not in use. Our proposed system aims to create an advancement in education systems by face detection and recognition-based attendance system, speech-to-text transcription for digital course material, and to save energy by automatically switching off the smart board when not in use.

Keywords Eigen faces techniques · Google APIs · Open CV · AdaBoost · Raspberry Pi · Haar classifier · Hidden Markov model · Passive infrared sensor · Convolutional neural network

1 Introduction

1.1 A Subsection Sample

Quality education, the involvement of the students in the classroom physically as well as mentally, and energy saving have become a difficult task for educational institutes. Smart classroom service proves to be one of the prominent technologies in

S. Agrawal · A. Chandhok · P. Sasikumar
School of Electronics Engineering, Vellore Institute of Technology, Vellore, India
e-mail: sasikumar.p@vit.ac.in

S. Maheswari (✉)
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India
e-mail: maheswari.s@vit.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_74

the IoT era that can lead us to solutions for the listed problems. The main objective is to change the classroom into being a more intelligent, interconnected, and remotely controllable one which minimizes the wastage of energy. For smart attendance, the camera is fixed at the entrance of the classroom which will capture the image of the student. Machine learning algorithms like AdaBoost, Haar Classifier, histogram of oriented gradients, and lower binary pattern histogram face recognizer have been used and the attendance would be marked accordingly. This model converts class lectures into text using Google Speech Recognition that automatically recognizes the words that are spoken and converts them in text format. This will be saved in the form of a text file which could be made available to students through the web page. Passive infrared sensor (PIR) senses the infrared radiation transmitted from humans in its field of view and when no human is detected, it would switch off the smart board.

2 Literature Survey

With the advancement in technology, there exist many ways of capturing students' attendance. In the education sector, the way people access information has been revolutionized at an extraordinary rate owing to the extensive use of smartphones. There are various ways in which attendance capturing is implemented. Some of the existing works for a fast and smart attendance system is by the means of Bluetooth low-energy sensors which can be clung to the identification card of each student which contains a unique string which is further associated with a mobile application that is made accessible to the faculty that gathers the data and marks the attendance accordingly [1]. Although it can facilitate faculties to take attendance, it failed to be the best solution because of the short-range of Bluetooth sensors. Another study suggested an attendance system which required a QR code scan with the use of a Web camera affixed to a computer. The gathered data in the form of image was saved in a database for the purpose of retrieval [2]. A similar implementation suggested a prototype for the implementation of an attendance system for faculties as well as students that is fully automated by using the concept of face detection and recognition technique. For this purpose, various approaches have been used including CNN, Eigenface values as well as principal component analysis. After this, a comparison was made for the faces detected and the faces of the students already present in the database [3].

Currently, the electricity supply is not able to meet the exponentially increasing demands of the mass. For our future generation, we need to save energy as a slogan suggests "One unit saved is one unit generated." With continuous research and technical improvements, researchers have been focused on increasing luxury in life along with conservation in energy. The automation of home and offices was very well received by people from all over the world. In 2011, a home automation design was suggested where an Arduino BT board was used along with relays for connecting all appliances to the I/O ports of the board. A cell phone was used to send signals

to the Arduino using Bluetooth technology which further switched on/off the appliances [4]. Though it is a low-cost, flexible, and wireless solution, it made charged cell phones a necessity and required manual signaling. Later on, researchers found infrared energy emitted from the objects can be sensed by the PIR sensor which can be used to control electronic and electrical devices. Using this concept, a prototype is built for a system that facilitates energy conservation with the help of a cascade of PIR sensors to control fans and lights present in the class and turn it off in the absence of students [5]. It is further suggested that fan speed can be controlled automatically according to our requirements. The LM35 temperature sensor senses the room temperature and converts it into an electrical signal, which is sent to the microcontroller which then drives the transistor to control the speed of the fan [6].

For lecture transcription, Google Speech API is used. Google APIs have been implemented in many existing works like a chatbot. In this, a scalable software solution was built to collect and process speech to text. This model was based on the Python language and Web framework Flask. According to the results and after successful analysis of the collected dataset of texts and audios, a custom model was developed based on Keras library using a recurrent neural network (RNN) for training [7]. Yet another application for Google Speech Recognition API is in healthcare [8]. This work demonstrated that in services occurring during the time of operation, voice recognition technology is used to permit perioperative staff to monitor the status of workflow with the help of mobile-based technology. By using some of the classifiers like maximum entropy, bag-of-words, and support vector machine, this work further improved the recognition capability of Google's voice. Yet another study developed a software named SPEECH-TO-TEXT that allows the user to control computer functions and repeats text by voice input. Hidden Markov model is used in building a framework for each letter. Mel frequency Cepstral coefficients help in feature extraction. For the training of the dataset, vector quantization is being used, and testing of the dataset is done with the help of the Viterbi algorithm [9].

2.1 Block Diagram

Here, Fig. 1 shows the block diagram of the project which mainly consists of three functionalities, namely attendance using face detection and recognition, lecture transcription, automation of smart board.

2.2 Technical Specification

The hardware and software used in this project include "Python" coding for face detection, recognition, and automation of smart board using PIR sensors. Google Speech Recognition API is used for speech-to-text conversion.

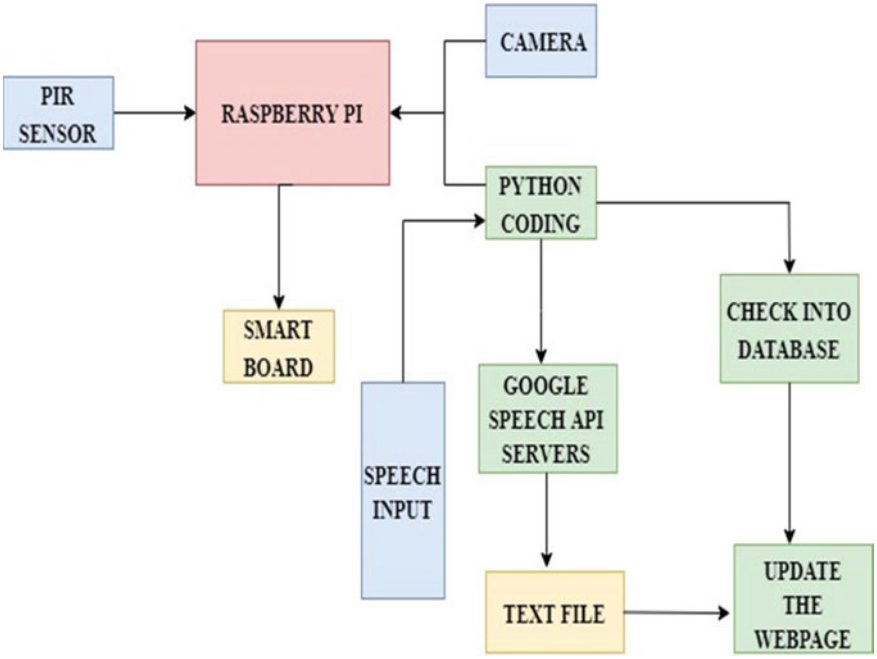


Fig.1 Block diagram of smart classroom model

2.2.1 Hardware

The hardware used here are: Raspberry Pi, Raspberry Pi Camera Module, PIR sensor, and Bluetooth microphone.

2.2.2 Software

The following software components are used on the development end:

Operating System (OS) used: Windows 10, Raspbian, Language used: Python 2.7, Python 3.5, Platform used: OpenCV (Linux Library).

3 Algorithms

Python language has been used for writing the algorithms. Our proposed model includes algorithms such as Haar classifier, histogram of oriented gradients (HOG), and lower binary pattern histogram (LBPH) face recognizer.

3.1 Haar Classifier

This OpenCV feature is used for the purpose of identifying the objects present in an image which is based on the concept of features and to train the classifier using a number of positive and negative images [10]. The images that we want for our classifier are termed as positive images and rest everything else is termed as negative images. This algorithm mainly includes four parts: the selection of Haar features, creation of integral images, AdaBoost training, and finally concatenation of the classifiers. Firstly, in a detection window, the adjacent regions which are rectangular are taken into account for Haar features selection. The intensities of the pixel in each of these regions are added, and then, the difference between these sums is computed.

3.2 Histogram of Oriented Gradients (HOG)

The histogram of oriented gradients is one of the feature descriptor algorithms used for the purpose of detection of any object in the field of image processing which is a subset of computer vision [11].

The execution of the HOG descriptor algorithm is given as:

1. The image is divided into small regions that are connected to each other known as cells. For each of this cell, a histogram is computed of edge orientations or gradient directions for the pixels which are present within these cells.
2. Each cell is then discretized according to the gradient orientation into angular bins.
3. Each cell's pixel donates a gradient whose weight is computed automatically to its corresponding angular bin.
4. For spatial regions called blocks, groups of adjacent cells are taken into consideration. The formation of blocks using the grouping of cells forms the foundation for the standardization of histograms.
5. The representation of the block histogram is done by the standardized group of histograms, and the collection of such block histograms constitutes the descriptor.

Magnitude of the gradient is given by the formula:

$$G = \sqrt{x^2 + y^2} \tag{1}$$

Direction of the gradient is given by the formula:

$$\theta = \arctan(x/y) \tag{2}$$

where x-gradient lies on a horizontal line, and y-gradient lies on a vertical line (Fig. 2).

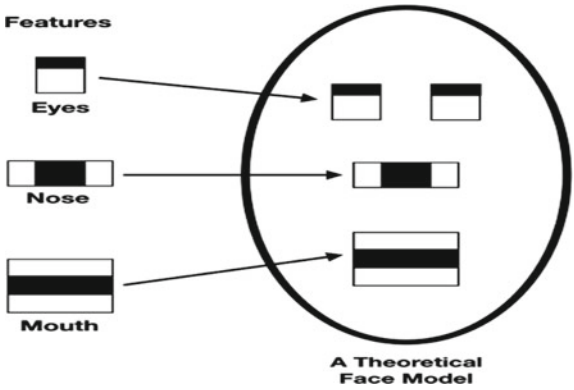


Fig. 2 Haar feature collection (Retrieved from www.becominghuman.ai)

3.3 Lower Binary Pattern Histogram (LBPH) Face Recognizer

This algorithm is part of the OpenCV library and primarily depends upon the local binary operator. This is mainly used to display the local features which are present in the images. The value of the pixels of an image is replaced with decimal numbers by LBP operators [12, 13]. Using the given image, the four-step process for the calculation of the LBP descriptor is given as:

1. In any image, I, for every pixel (a, b), N neighboring pixels are chosen with an R radius.
2. The difference in the intensity of the current pixel (a, b) with that of N numbers of neighboring pixels is calculated.
3. The intensity difference threshold forms a bit vector by assigning the value 0 to all the negative differences and value 1 to all the positive differences.
4. The intensity value at (a, b) is replaced by the decimal value by converting the N-bit vector value to its equivalent decimal number.

For every pixel, the LBP descriptor is defined as:

$$\mathrm{LBP}\,(N, R) = \sum_{n=0}^{N-1} f\left(i_p - i_c\right)2^n \tag{3}$$

where i_p represents neighboring pixel's intensity, and i_c represents the current pixel's intensity. All the neighboring pixels are denoted by N that are chosen at a R radius.

4 Results and Outputs

The entire smart classroom model worked as expected and gave the results with greater accuracy and fulfilled all the objectives that have been stated; used training dataset is shown in Fig. 3. The smart attendance system using face identification and recognition gave the correct output and successfully updated the same in the Web portal as shown in Fig. 4. The digital transcription functionality can convert the class lecture into the text format with the maximum accuracy possible and the same is uploaded on the Web page as well. Automation of the smart board is also performed successfully by setting a delay of 1 min so that the faculty can take a round of the class without the smart board turning off. The entire model successfully converted the classroom environment into a SMART one.

Fig. 3 Training dataset

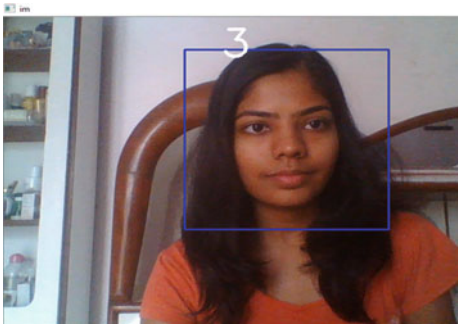


Fig. 4 Output for facial recognition-based attendance system

5 Conclusion

In the project presented, a smart classroom has been implemented wherein facial recognition and speech-to-text transcription and automation have been made use of. Facial recognition-based attendance developed using Python libraries and machine learning algorithms helped in the utilization of work time efficiently. Lecture transcription successfully converted lecture notes into digital format, enabling students to not only gain more knowledge during the class but also providing remote access to the notes. The integration of these subsystems helped create a classroom-friendly environment with a nominal one-time cost. Our project aspires to be an advanced technology making a mark in the current smart classroom technologies.

6 Future Scope

The proposed smart classroom model can be further improvised by using high megapixel cameras and improved algorithms for better accuracy when students enter in a group rather than a queue. A biometric scan can be added to the facial recognition-based system for additional security. For lecture transcription, algorithms can be added which can put punctuations and for sentences from the text based on previous training.

References

1. Apoorv R, Mathur P (2016) Smart attendance management using bluetooth low energy and android. In: 2016 IEEE region 10 conference (TENCON). IEEE, pp 1048–1052
2. Hendry MRM, Rahman MNA, Seyal AH (2017) Smart attendance system applying QR code. In: 12th international conference on latest trends in engineering and technology (ICLTET'2017), pp 22–24
3. Sawhney S, Kacker K, Jain S, Singh SN, Garg R (2019) Real-time smart attendance system using face recognition techniques. In: 2019 9th international conference on cloud computing, data science & engineering (Confluence). IEEE, pp 522–525
4. Moreira RDS, Ebecken NFF, Alves AS, Livernet F, Campillo-Navetti A (2014) A survey on video detection and tracking of maritime vessels. Int J Recent Res Appl Stud 20(1)
5. Twumasi C, Dotche KA, Banuenumah W, Sekyere F (2017) Energy saving system using a PIR sensor for classroom monitoring. In: 2017 IEEE PES power Africa. IEEE, pp 347–351
6. Baligar S, Joshi S, Mudhole S, Jadhav S, Jambotkar C (2019) Temperature based speed control of fan using Arduino
7. Ranchal R, Doughty T, Guo Y, Bain K, Martin H, Robinson J, Duerstock B (2013) Using speech recognition for real-time captioning and lecture transcription in the classroom. IEEE Trans Learn Technol 6:299–311. https://doi.org/10.1109/TLT.2013.21
8. Uddin M, Huynh N, Vidal JM, Taaffe K, Fredendall LD, Greenstein JS (2015) Evaluation of Google's voice recognition and sentence classification for health care applications
9. Das P, Acharjee K, Das P, Prasad V (2015) Voice recognition system: speech-to-text. J Appl Fundam Sci 1:2395–5562

10. 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. https://doi.org/10.1109/CVPR.2001.990517
11. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol 1. IEEE, pp 886–893
12. Deeba F, Memon H, Dharejo F, Ahmed A, Ghaffar A (2019) LBPH-based enhanced real-time face recognition. Int J Adv Comput Sci Appl 10. https://doi.org/10.14569/IJACSA.2019.0100535
13. Rahim Md, Hossain Md, Wahid T, Azam Md (2013) Face recognition using local binary patterns (LBP)
14. Potadar MP, Marathe VV, Khose AS, Kotkar LA (2015) Biometric attendance recording and communication system. Int J Innov Eng Technol (IJIET) 5(2):230–234

RADAR and Camera Sensor Data Fusion

Venkatesh Mane, Ashwin R. Kubasadgoudar, P. Nikita, and Nalini C. Iyer

Abstract As demand for vehicle automation has expanded in the last few years, it has become imperative to more precisely recognize the position and speed of surrounding vehicles. Object detection has been recognized as an important feature of the advanced driving assistance system (ADAS). This also ensures the safety of vehicles and prevents accidents caused by the negligence of humans. Object detection with sensor data fusion has proved to be very effective. Obstacles can be detected and labeled with the help of RADAR, LIDAR, and camera. Every sensor has advantages and limitations. Limitations of one sensor can be overcome by another sensor. Sensors such as LIDAR, RADAR, and camera are used together in order to obtain optimum results contributing to better object detection in autonomous systems. The paper describes the fusion of data acquired by the two sensors RADAR (AWR1642BOOST) and a two-dimensional camera (LOGITECH C170). RADAR can achieve better results in distance calculation than camera, whereas camera can achieve better results in angle compared to RADAR. Similarly, RADAR works efficiently in poor weather conditions and lighting, whereas camera may not provide accurate results. The data acquired by both the sensors are fused in order to obtain better object detection and ensure accurate calculation of parameters of the object detected. Region of interest detection and Haar Cascade algorithms are implemented to get satisfactory results and has been implemented in real time.

Keywords RADAR · Sensor fusion · CAMERA · AWR1642 · ADAS · LIDAR · Camera · Object detection

V. Mane (✉) · A. R. Kubasadgoudar · P. Nikita · N. C. Iyer
KLE Technological University, Hubli, India
e-mail: mane@kletech.ac.in

A. R. Kubasadgoudar
e-mail: ashwinrk@kletech.ac.in

P. Nikita
e-mail: nikhita_patil@kletech.ac.in

N. C. Iyer
e-mail: nalinic@kletech.ac.in

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2_75

Table 1 Performance of RADAR and camera for different conditions

Condition	RADAR	Camera	Fusion
Obstacle	Excellent	Fair	Excellent
Distance	Excellent	Fair	Excellent
Weather	Excellent	Limited	Excellent
Lighting	Excellent	Limited	Excellent
Boundary	Limited	Excellent	Excellent
Angle	Fair	Excellent	Excellent

1 Introduction

Vehicle automation is an essential topic in today's age. Object detection is an integral feature of the advanced driver assistance system (ADAS); this not only eases the driving experience for the driver but also provides security. Many accidents in the world occur due to the negligence of drivers, which causes them to hit or crash into objects around them. Object detection helps detect and label objects while driving, also indicates the orientation of the object with respect to the vehicle [1, 2]. This can be achieved by fusing data between two or more sensors, ensuring accurate and precise results and parameter calculation.

Table 1 represents the performance of RADAR and camera in different conditions.

The RADAR is an excellent source for object detection as compared to the camera, which performs a pretty good job at object detection [3, 4]. RADAR performs excellently for distance calculation when compared to the camera.

2 Proposed Design

Camera and RADAR, the two sensors being interfaced, have properties that can be used to their best abilities to obtain superlative results. RADAR provides accurate measurement for obstacle distance and velocity but fails to identify boundary detection, extraction of features, and the angle which can be obtained with the camera. Considering the complementary properties of both the camera and RADAR, the sensors are fused in order to get the optimum results. Object detection can be realized by analyzing different algorithms that can be used in the project, "RADAR and Camera Sensor Data Fusion." Pre-processing of data being an important step is realized by different methods. In this chapter, an appropriate system is proposed after looking into various research elements about object detection. The basic working and the functional block diagram of the system are discussed.

(A) Functional block diagram

Figure 1 proposed block diagram comprises the smaller functional blocks and their individual functionality, which together are interconnected to form the fully functional system design.

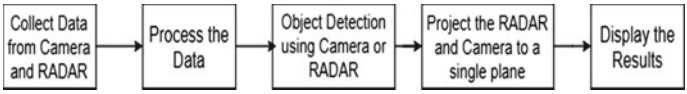
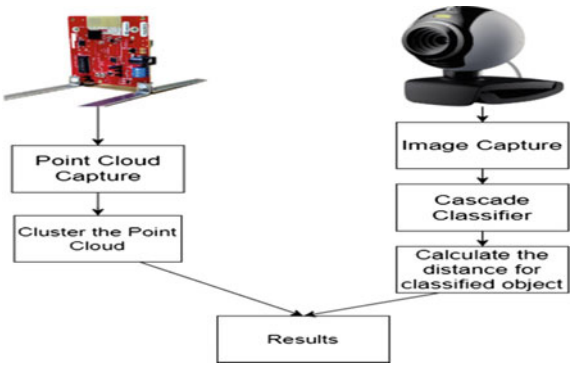


Fig. 1 Functional block diagram

Fig. 2 Functional block diagram



The data is obtained from both sensors and processed with their respective algorithms. Both the sensor data are fused to a single plane to obtain more precise object detection.

(B) Proposed System Design

An optimal solution is selected based on its working and ease of implementation, as shown in Fig. 2. Point cloud data from the RADAR and image data from the camera are captured. The point cloud obtained is clustered; this happens to be a significant step as the load on the program is reduced [5, 6]. The image captured goes through a Haar Cascade. The velocity and distance from the RADAR are compared with that of the camera giving the final compiled and compared results.

3 RADAR and Camera Specifications

The RADAR used is the Texas instrument's AWR1642BOOST mmWave RADAR. This RADAR is an FMCW (Frequency Modulated Continuous Wave) RADAR that can operate at a frequency range of 76–81 GHz. It has four receive and 2 transmit patch antennas; these can be individually selected based on the application required. The configuration for the range, range resolution, and number of antennas to be used should be made before setting it up. This can be done using mmWave Demo Visualizer and Uniflash tool provided by the Texas instruments.

The RADAR configuration is as follows (Table 2):

Table 2 RADAR configuration

Base frequency	77 GHz
Sweep time	100 ms
Antenna configuration	4 Rx, 2 Tx
Max range	10 m
Range resolution	0.044 m
Azimuth resolution	15°
mmWave SDK	v1.2

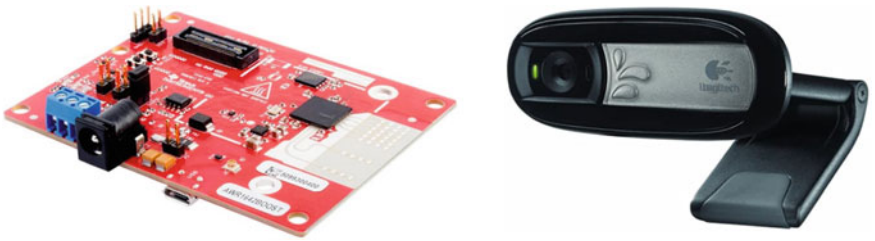


Fig. 3 AWR1642BOOST and Logitech C170 webcam

The camera used is the Logitech C170 webcam. The resolution of the camera is 11.3 MP, which captures an image in the dimensions of 640 × 480 pixels. The field of view (FOV) of the camera is 69° and has been experimentally verified. Figure 3 is the image of the Logitech C170 webcam being used.

4 Implementation Details

(A) Algorithms

Algorithms are utilized for RADAR and camera, which will be used for fusion further. The algorithms such as DBSCAN (density-based clustered algorithm) for clustering the points of RADAR and Haar Cascade algorithm for camera object detection are explained in this section. *DBSCAN (Density-Based Clustered Algorithm)*: The outcome of AWR1642BOOST for detecting obstacles provides point cloud data. These points provide a wide range of sensitive data, which is challenging to compute. Clustering of these points will help for more accurate computation of RADAR data. It tends to give multiple points for a single object because of high range resolution. DBSCAN is a density-based clustered algorithm is a sliding window-based algorithm that attempts to find dense areas of data points (Fig. 4).

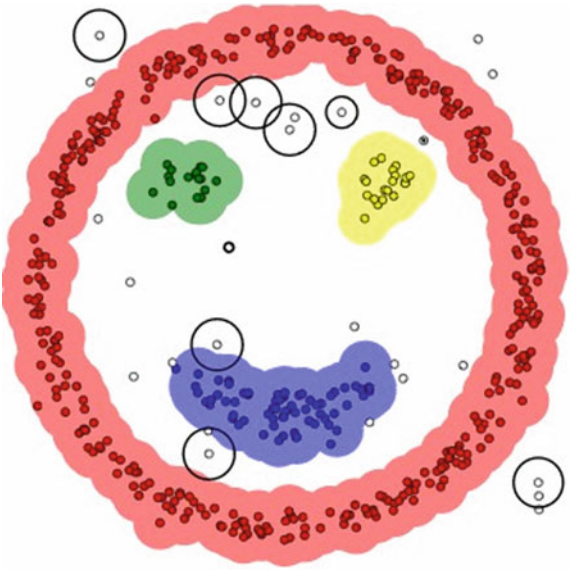


Fig. 4 DBSCAN clustering algorithm

Application of DBSCAN in the proposed design:

- For this first point in the new cluster, the points within its epsilon distance neighborhood also become part of the same cluster. This procedure of making all points in the epsilon neighborhood belong to the same cluster is then repeated for all of the new points that have been just added to the cluster group.
- After completing the current cluster, retrieving new unvisited points is done and processed, which then leads to the discovery of a further cluster or noise.

(1) *Camera Calibration*: The camera has two kinds of parameters, namely extrinsic and intrinsic parameters.

1. Intrinsic parameters estimates single-instrument operational parameters. Hence, methods for finding intrinsic parameters for different types of sensors do not share many similarities. The inherent criteria for a sensor include calculating the focal length, the reference focus, or the optical center, skew factors and criteria of lens distortion. In order to obtain optimum results, fusion of the data obtained from both camera and RADAR is done. The fusion of data aids in the decision-making that is required for the fusion.
2. Extrinsic parameters estimate the spatial transformation of the sensor coordinates to the other sensors or frame of reference. This is also recognized as calculating poses and tracking sensors. Extrinsic calibration parameterization, that is, homogeneous transformation, can always be expressed in the same way, regardless of the sensors involved.

In order to obtain optimum results, fusion of the data obtained from both camera and RADAR is done. The fusion of data aids in the decision-making that is required for the fusion. For fusion to occur, both camera and RADAR need to be in the same measurement system so that comparison and communication can effortlessly take place.

For the camera to communicate with the RADAR, converting it into real-world coordinates is done by calibrating the camera.

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} fx & 0 & cx \\ 0 & fy & cy \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r11 & r12 & r13 & t1 \\ r21 & r22 & r23 & t2 \\ r31 & r32 & r33 & t3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Camera parameters are represented in a 4-by-3 matrix called the camera matrix. This matrix maps the 3D world scene into the image plane. The calibration calculation calculates the camera matrix utilizing the extrinsic and intrinsic parameters. The extrinsic parameters speak to the area of the camera within the 3-D scene. The intrinsic parameters represent the optical center and central length of the camera.

(B) Distance Calculation

(1) *Camera*: The camera uses its focal length to estimate the distance of the object from it. The focal length of the Logitech C170 webcam is 2.9 mm and can be found out by checking the properties of the camera after clicking a picture with this camera. The distance of the object can be calculated using an equation.

$$R = \frac{f \times \text{height real} \times \text{height image pixels}}{\text{height object pixels} \times \text{height sensor}} \tag{1}$$

In Eq. 1, the height and width of the object are obtained by applying the Haar Cascade algorithm, which is then substituted in the equation to obtain the distance of the object from the camera. This is used to find the region of interest and also used to match the point cloud readings from the RADAR.

(2) *RADAR*: RADAR uses the principles of the Doppler effect to calculate the coordinates of the object. FMCW waves are the transmission waves that have been used in AWR1642BOOST automotive RADAR. Doppler effect is the apparent shift in the frequency of the transmitted wave when there is relative motion between the target and the sensor. The Dechirp process is used to find the apparent shift in the frequency of the signal. In Eq. 2, X_1 is X and X_2 is the RefX, and X_{out} is the Dechirp signal in the block diagram.

$$X_{\text{out}} = X_1 + X_2 \tag{2}$$

where

$$X_1 = \sin[(\omega_1 t + (\emptyset_1))] \tag{3}$$

$$X_2 = \sin[(\omega_2)t + (\emptyset_2)] \tag{4}$$

that is

$$X_{\text{out}} = \sin[(\omega_1 + \omega_2)t + (\emptyset_1 - \emptyset_2)] \tag{5}$$

Using the shift in the frequency, the range of the target can be obtained with Eq. 7

$$\Delta f = k \cdot \Delta t \tag{6}$$

$$R = \frac{c}{2} * \frac{\Delta f}{k} \tag{7}$$

where Δf  Apparent shift in freq, k  Slope of the graph, $\Delta t \rightarrow$ time.

(C) Velocity Calculation

The video frame is recorded, and the width of the road in pixels is calculated digitally. Now that the width of the road in meters is obtained from the real world and in pixels from the video frame that has been recorded, to map the distances between these two worlds, calculated pixels per meter are used to divide the distance of road in pixels to meters. dpixels gives the pixel distance traveled by the vehicle in one frame of our video processing. To estimate speed in any standard unit first, there is a need to convert dpixels to dmeters.

$$\text{speed} = \text{dmeters} * \text{fps} \tag{8}$$

dmeters is the distance traveled in one frame. Calculated the average fps during video processing. So, to get the speed in m/s

$$\text{speed} = \text{dmeters} * \text{fps} * 3 : 6 \tag{9}$$

(D) Angle Calculation

Angle calculation is done by the camera algorithm. Since a digital image is composed of both rows and columns, the angle can be calculated using a column number. The below equation states the calculation for the angle using the camera. In the equation,

x represents the column number of the detected object, and the denominator is equal to 9 as the number of columns divided by the field of view of the camera. 320 is the middle column number. The maximum field of view is 69°, which is equal to 34.5° on either side. Hence, the maximum angle that can be measured is 34.5°. When a person stands directly in front of the camera, then the angle is 0°. The angle varies as the person moves on either side of the camera from + 34.5 to −34.5. The angle is 20° when the person stands on either side of the camera. In Fig. 5, the object/person is standing right in front of the camera. Hence, the angle is taken as 0°.

$$\text{angle} = \frac{x - 320}{9} \tag{10}$$

In Fig. 6, the object/person is 20° to the right or to the left of the camera. Hence, the angle is taken as 20°.

(E) Haar Cascade Classifier

Haar Cascade classification technique is based on Haar Wavelet technique to analyze pixels in the image into squares by function and Viola–Jones detection algorithm. This method is used to detect the face of the person from the image captured from the camera. Haar Cascade is a machine learning-based approach where many positive and negative images are used to train the classifier. It is collecting Haar characteristics. A Haar function considers neighboring rectangular regions in a detection window

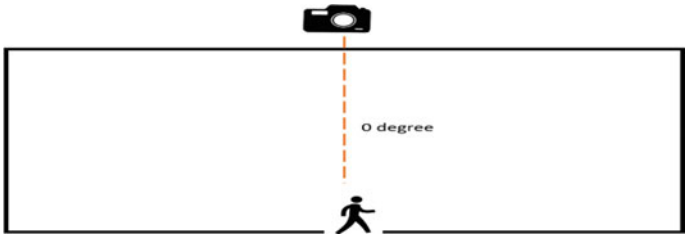


Fig. 5 0° angle calculation using camera

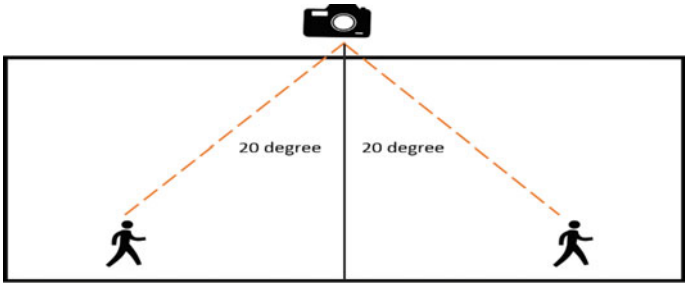


Fig. 6 20° angle calculation using camera

at a given position, sums up the intensities of pixels in each area, and measures the difference between those quantities. A window of the target size is moved over the input image during the detection phase, and the image and Haar features are calculated for each subsection. The learned threshold is then used to compare which separates non-objects from objects.

5 Results and Discussions

For real-time object detection, sensors and fusion of RADAR and camera using the methodology above explained are detailed in this forthcoming chapter.

5.1 Result Analysis

A person is detected by the camera, and the diagonal distance is obtained with the algorithm. The center of the frame is the reference plane. The distance from the reference plane to the person is calculated, then the diagonal distance is obtained. The result can be seen in Fig. 7.

RADAR data represents the detected objects as points. These points provide a wide range of sensitive data, which is challenging to compute. Clustering of these points will help for more accurate computation of RADAR data. Figure 8 represents the data before applying the DBSCAN algorithm and after applying DBSCAN algorithm. Figure 9 describes the image obtained by the camera sensor with processed algorithm resulting with distance and the camera angle. The distance from the camera sensor to the obstacle or pedestrian is detected by the camera and is denoted in the figure. Angle is determined considering the center of the frame as zero, and it can be noticed that

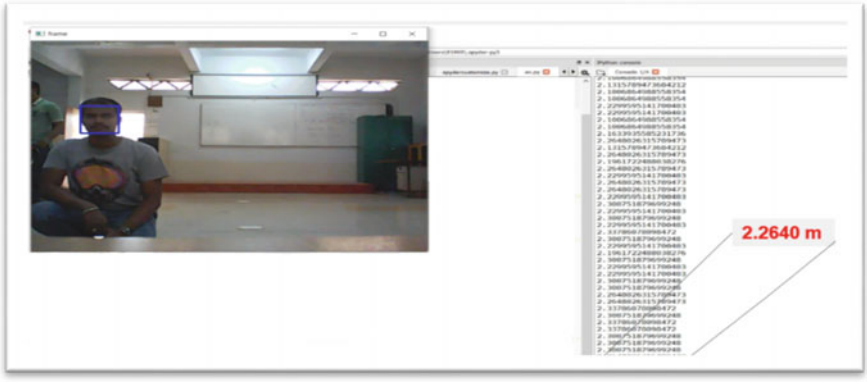


Fig. 7 Results obtained from camera

Fig. 8 RADAR detected obstacles before clustering and after DBSCAN algorithm

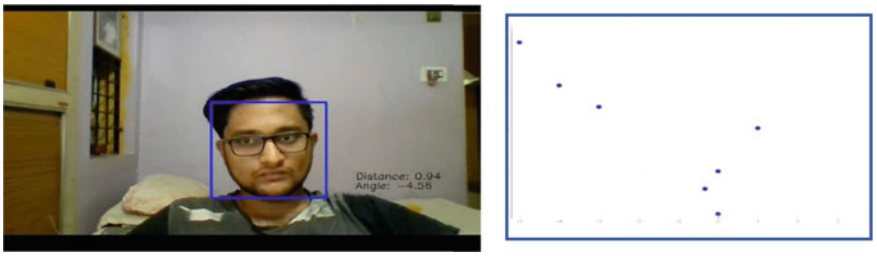


Fig. 9 Fusion camera frame and fusion RADAR frame

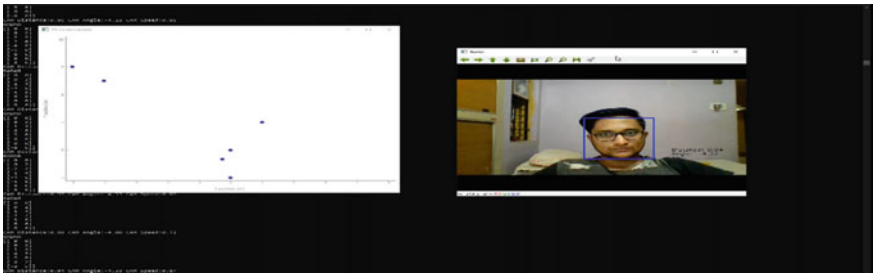


Fig. 10 Console snapshot

the obstacle is 4.56° right of the camera. The bounding box in the image is the output of the Haar Cascade algorithm as camera performs better in plotting the boundary box than RADAR. Figure 9 represents RADAR obtained points after DBSCAN. The result is corresponding data of the image captured by the camera where the distance is determined by the RADAR as it performs well for distance calculation than camera. In the plot, it is observed that the distance obtained by RADAR for a particular obstacle is a 1 m distance approximately which is denoted in the camera frame as distance: 0.94.

Figure 10 is the snapshot of the fused data from both the sensor with RADAR points in the command prompt.

References

1. Yu Z, Bai J, Chen S, Huang L et al (2018) Camera-radar data fusion for target detection via Kalman filter and Bayesian estimation. SAE Technical Paper 2018-01-1608
2. Kim K, Lee C, Pae D, Lim M (2017) Sensor fusion for vehicle tracking with camera and radar sensor. In: 2017 17th international conference on control, automation and systems (ICCAS). Jeju, pp 1075–1077. https://doi.org/10.23919/ICCAS.2017.8204375
3. Nobis F, Geisslinger M, Weber M, Betz J, Lienkamp M (2019) A deep learning-based radar and camera sensor fusion architecture for object detection. In: 2019 sensor data fusion: trends, solutions, applications (SDF). Bonn, Germany, pp 1–7. https://doi.org/10.1109/SDF.2019.8916629
4. Burlet J, Dalla Fontana M (2012) Robust and efficient multi-object detection and tracking for vehicle perception systems using radar and camera sensor fusion. In: IET and ITS conference on road transport information and control (RTIC 2012). London, pp 1–6. https://doi.org/10.1049/cp.2012.1553
5. Coates A, Ng AY (2010) Multi-camera object detection for robotics. In: 2010 IEEE international conference on robotics and automation, Anchorage, AK, pp 412–419. https://doi.org/10.1109/ROBOT.2010.5509644; bibitemb6 Zewge NS, Kim Y, Kim J, Kim J (2019) Millimeter-wave radar and RGB-D camera sensor fusion for real-time people detection and tracking. In: 2019 7th international conference on robot intelligence technology and applications (RiTA). Daejeon, Korea (South), pp 93–98
6. Ji Z, Prokhorov D (2008) Radar-vision fusion for object classification. In: 2008 11th international conference on information fusion. Cologne, pp 1–7

Author Index

A. Joshi et al. (eds.), *Information and Communication Technology for Competitive Strategies (ICTCS 2021)*, Lecture Notes in Networks and Systems 400,
https://doi.org/10.1007/978-981-19-0095-2