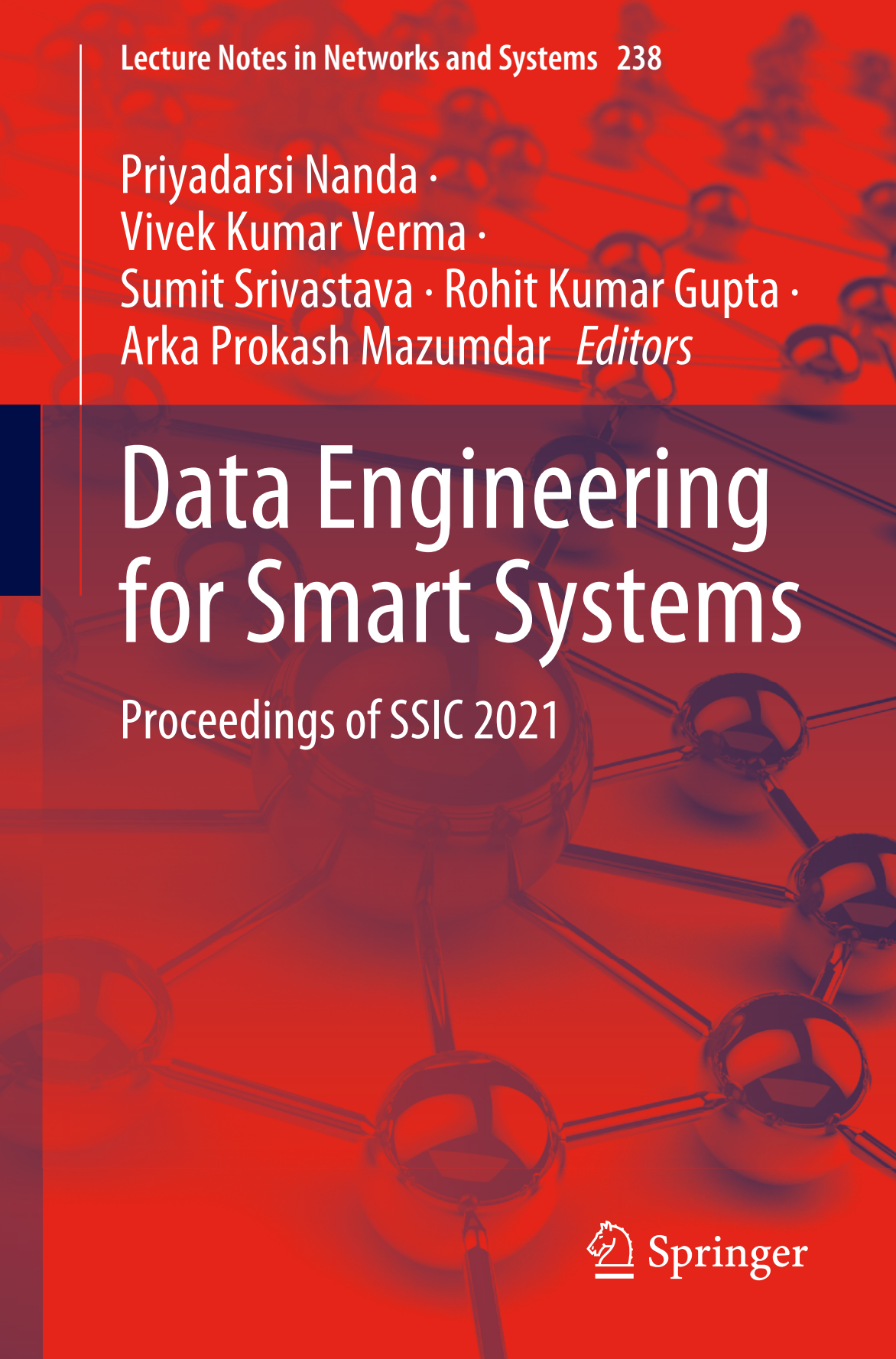
Lecture Notes in Networks and Systems 238

Priyadarsi Nanda ·
Vivek Kumar Verma ·
Sumit Srivastava · Rohit Kumar Gupta ·
Arka Prokash Mazumdar *Editors*

Data Engineering for Smart Systems

Proceedings of SSIC 2021



Springer

Lecture Notes in Networks and Systems

Volume 238

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

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

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

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

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

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

More information about this series at http://www.springer.com/series/15179

Priyadarsi Nanda · Vivek Kumar Verma ·
Sumit Srivastava · Rohit Kumar Gupta ·
Arka Prokash Mazumdar
Editors

Data Engineering for Smart Systems

Proceedings of SSIC 2021

Springer

Editors
Priyadarsi Nanda
School of Electrical and Data Engineering
University of Technology Sydney
Sydney, NSW, Australia

Sumit Srivastava
Manipal University Jaipur
Jaipur, Rajasthan, India

Arka Prokash Mazumdar
Department of Computer Science
and Engineering
Malviya National Institute of Technology
Jaipur, Rajasthan, India

Vivek Kumar Verma
Manipal University Jaipur
Jaipur, Rajasthan, India

Rohit Kumar Gupta
Manipal University Jaipur
Jaipur, Rajasthan, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-16-2640-1 ISBN 978-981-16-2641-8 (eBook)
https://doi.org/10.1007/978-981-16-2641-8

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

This book is one-pot solution for authors to showcase their work among the research communities, application warehouses and the public–private sectors. This series is an opportunity to gather research scientific works related to data engineering concept in the context of computational intelligence consisted of interaction between smart devices, smart environments and smart interactions, as well as information technology support for such areas. Data from mentioned areas also need to be stored (after their gathering) in intelligent database systems and to be processed (using smart and intelligent approach). The aim of this series is to make available a platform for the publication of books on all aspects of single- and multi-disciplinary research on these themes to make the latest results available in a readily accessible form. The innovations provide the latest software tools to be showcased in the book series through publications. The high-quality content with broad range of the topics pertaining to the book series will be peer-reviewed and get published on suitable recommendations.

This book will provide state of the art to research scholars, scientists, industry learners and postgraduates from the various domains of engineering and related fields, as this incorporates data science and the latest innovations in the field of engineering with their paradigms and methods that employ knowledge and intelligence in the research community. This book comprises its scope ranging from data science for smart systems which are having inbuilt capabilities of handling the research challenges and problems related to energy-aware sensors, smart city projects, wearable devices, smart healthcare solutions, smart e-learning initiatives and social implications of IoT. Further, it extends its coverage to the different computational aspects involved in various domains of engineering such as complex security solutions for data engineering, communication networks, data analytics, machine learning, integrating IoT data with external data sources and data science approaches for smart systems.

Secondarily, this book provides the technological solutions to non-engineering and sciences domain as it contains the fundamental innovation in the field of engineering which turns to be a real solution for their problems. Also, it includes the

paradigms which support industries for the development of solutions in favor of society and make everyone's life easier.

Sydney, Australia — Priyadarsi Nanda
Jaipur, India — Vivek Kumar Verma
Jaipur, India — Sumit Srivastava
Jaipur, India — Rohit Kumar Gupta
Jaipur, India — Arka Prokash Mazumdar

Contents

Using Machine Learning, Image Processing and Neural Networks to Sense Bullying in K-12 Schools: Enhanced 1
Lalit Kumar, Palash Goyal, Karan Malik, Rishav Kumar, and Dhruv Shrivastav

Feature-Based Comparative Study of Machine Learning Algorithms for Credibility Analysis of Online Social Media Content 13
Utkarsh Sharma and Shishir Kumar

Smart Support System for Navigation of Visually Challenged Person Using IOT 27
Tuhin Utsab Paul and Aninda Ghosh

Identity-Based Video Summarization 37
Soummya Kulkarni, Darshana Bhagit, and Masooda Modak

Security Testing for Blockchain Enabled IoT System 45
A. B. Yugakiruthika and A. Malini

Two-Dimensional Software Reliability Model with Considering the Uncertainty in Operating Environment and Predictive Analysis 57
Ramgopal Dhaka, Bhoopendra Pachauri, and Anamika Jain

Object Recognition in a Cluttered Scene 71
Rashmee Shrestha, Mandeep Kaur, Nitin Rakesh, and Parma Nand

Breast Cancer Prediction on BreakHis Dataset Using Deep CNN and Transfer Learning Model 77
Pinky Agarwal, Anju Yadav, and Pratistha Mathur

A Comprehensive Tool Survey for Blockchain to IoT Applications 89
A. B. Yugakiruthika and A. Malini

Hybrid Ensemble for Fake News Detection: An Attempt 101
Lovedeep Singh

Detection of Abnormal Activity at College Entrance Through Video Surveillance 109
Lalit Damahe, Saurabh Diwe, Shailesh Kamble, Sandeep Kakde, and Praful Barekar

Audio Peripheral Volume Automation Based on the Surrounding Environment and Individual Human Listening Traits 123
Adit Doshi, Helly Patel, Rikin Patel, Brijesh Satasiya, Muskan Kapadia, and Nirali Nanavati

A Survey: Accretion in Linguistic Classification of Indian Languages 133
Dipjayaben Patel

The Positive Electronic Word of Mouth: A Research Based on the Relational Mediator Meta-Analytic Framework in Electronic Marketplace 147
Bui Thanh Khoa

A Review: Web Content Mining Techniques 159
Priyanka Shah and Hardik B. Pandit

GPS-Free Localization in Vehicular Networks Using Directional Antennas 173
Parveen, Sushil Kumar, and Rishipal Singh

An Improved Scheme in AODV Routing Protocol for Enhancement of QoS in MANET 183
Amit Kumar Bairwa and Sandeep Joshi

Knowledge Management Framework for Sustainability and Resilience in Next-Gen e-Governance 191
Iqbal Hasan and Sam Rizvi

Enriching WordNet with Subject Specific Out of Vocabulary Terms Using Existing Ontology 205
Kanika, Shampa Chakraverty, Pinaki Chakraborty, Aditya Aggarwal, Manan Madan, and Gaurav Gupta

Automatic Detection of Grape, Potato and Strawberry Leaf Diseases Using CNN and Image Processing 213
Md. Tariqul Islam and Abdur Nur Tusher

Sentiment Analysis on Global Warming Tweets Using Naïve Bayes and RNN 225
Deborah T. Joy, Vikas Thada, and Utpal Srivastava

Novel and Prevalent Techniques for Resolving Control Hazard 235
Amit Pandey, Abdella K. Mohammed, Rajesh Kumar, and Deepak Sinwar

Sound Classification Using Residual Convolutional Network 245
Mahesh Jangid and Kabir Nagpal

Dimensionality Reduction Using Variational Autoencoders 255
Parv Dahiya and Sahil Garg

Computational Intelligence for Image Caption Generation 263
Sahil Garg and Parv Dahiya

An Empirical Statistical Analysis of COVID-19 Curve Through Newspaper Mining 271
Shriya Verma, Sonam Garg, Tanishq Chamoli, and Ankit Gupta

Wearable Monopole Antenna with 8-Shaped EBG for Biomedical Imaging 281
Regidi Suneetha and P. V. Sridevi

Speech Signal Processing for Identification of Under-Resourced Languages 291
Shweta Sinha

A Unique ECG Authentication System for Health Monitoring 299
Kusum Lata Jain, Meenakshi Nawal, and Shivani Gupta

Assamese Dialect Identification From Vowel Acoustics 313
Priyankoo Sarmah and Leena Dihingia

A Review on Metaheuristic Techniques in Automated Cryptanalysis of Classical Substitution Cipher 323
Ashish Jain, Prakash C. Sharma, Nirmal K. Gupta, and Santosh K. Vishwakarma

An Empirical Study of Different Techniques for the Improvement of Quality of Service in Cloud Computing 333
Chitra Sharma, Pradeep Kumar Tiwari, and Garima Agarwal

Contribution Analysis of Scope of SRGAN in the Medical Field 341
Moksh Kant, Sandeep Chaurasia, and Harish Sharma

Machination of Human Carpus 353
Sumit Bhardwaj, Bhuvidha Singh Tomar, Adarsh Ankur, and Punit Gupta

Building Machine Learning Application Using Oracle Analytics Cloud 361
Tarun Jain, Mahek Agarwal, Ashish Kumar, Vivek Kumar Verma, and Anju Yadav

An Empirical Analysis of Heart Disease Prediction Using Data Mining Techniques 377
Ashish Kumar, Sivapuram Sai Sanjith, Rajkishan Cherukuru, Vivek Kumar Verma, Tarun Jain, and Anju Yadav

2D Image to Standard Triangle Language (STL) 3D Image Conversion 391
Manoj K. Sharma and Ashish Malik

Improving Recommendation for Video Content Using Hyperparameter Tuning in Sparse Data Environment 401
Rohit Kumar Gupta, Vivek Kumar Verma, Ankit Mundra, Rohan Kapoor, and Shekhar Mishra

Importance and Uses of Telemedicine in Physiotherapeutic Healthcare System: A Scoping Systemic Review 411
Saurabh Kumar, Ankush Sharma, and Priyanka Rishi

Feature Exratction of PTTS System and Its Evaluation by Standard Statistical Method Mean Opinion Score 423
Sunil Nimbhore, Suhas Mache, and Sidhharth Mache

Computational Analysis of a Human–Robot Working Alliance Trust in Robot-Based Therapy 431
Azizi Ab Aziz and Wadhah A. Abdulhussain

Novel Intrusion Prevention and Detection Model in Wireless Sensor Network 443
Neha Singh and Deepali Virmani

Shadow Detection from Real Images and Removal Using Image Processing 451
Sumaya Akter Usrika and Abdus Sattar

A Study on Pulmonary Image Screening for the Detection of COVID-19 Using Convolutional Neural Networks 461
Shreyas Thakur, Yash Kasliwal, Taikhum Kothambawala, and Rahul Katarya

Analyzing Effects of Temperature, Humidity, and Urban Population in the Initial Outbreak of COVID19 Pandemic in India 469
Amit Pandey, Tucha Kedir, Rajesh Kumar, and Deepak Sinwar

FOFS: Firefly Optimization for Feature Selection to Predict Fault-Prone Software Modules 479
Somya Goyal

Deceptive Reviews Detection in E-Commerce Websites Using Machine Learning 489
Sparsh Kotriwal, Jaya Krishna Raguru, Siddhartha Saxena, and Devi Prasad Sharma

A Study on Buying Attitude on Facebook in the Digital Transformation Era: A Machine Learning Application 497
Bui Thanh Khoa, Ho Nhat Anh, Nguyen Minh Ly, and Nguyen Xuan Truong

Performance Evaluation of Speaker Identification in Language and Emotion Mismatch Conditions on Eastern and North Eastern Low Resource Languages of India 511
Joyanta Basu, Tapan Kumar Basu, and Swanirbhar Majumder

Autonomous Wheelchair for Physically Challenged 521
B. Kavyashree, B. S. Aishwarya, Mahima Manohar Varkhedi, S. Niharika, R. Amulya, and A. P. Kavya

Watermarking of Digital Image Based on Complex Number Theory 531
Nadia Afrin Ritu, Ahsin Abid, Al Amin Biswas, and M. Imdadul Islam

A Computer Vision Approach for Automated Cucumber Disease Recognition 543
Md. Abu Ishak Mahy, Salowa Binte Sohel, Joyanta Basak, Md. Jueal Mia, and Sourov Mazumder

An Automated Visa Prediction Technique for Higher Studies Using Machine Learning in the Context of Bangladesh 557
Asif Ahmmed, Tipu Sultan, Sk. Hasibul Islam Shad, Md. Jueal Mia, and Sourov Mazumder

COVID-19 Safe Guard: A Smart Mobile Application to Address Corona Pandemic 569
Soma Prathibha, K. L. Nirmal Raja, M. Shyamkumar, and M. Kirthiga

A Novel Mathematical Model to Represent the Hypothalamic Control on Water Balance 581
Divya Jangid and Saurabh Mukherjee

Importance of Deep Learning Models to Perform Segmentation on Medical Imaging Modalities 593
Preeti Sharma and Devershi Pallavi Bhatt

A Multi-component-Based Zero Trust Model to Mitigate the Threats in Internet of Medical Things 605
Y. Bevish Jinila, S. Prayla Shyry, and A. Christy

A Modified Cuckoo Search for the *n*-Queens Problem 615
Ashish Jain, Manoj K. Bohra, Manoj K. Sharma,
and Venkatesh G. Shankar

A Survey on Diabetic Retinopathy Detection Using Deep Learning 621
Deepak Mane, Namrata Londhe, Namita Patil, Omkar Patil,
and Prashant Vidhate

A Survey on Alzheimer's Disease Detection and Classification 639
D. T. Mane, Mehul Patel, Madhavi Sawant, Karina Maiyani,
and Divya Patil

An Approach for Graph Coloring Problem Using Grouping of Vertices 651
Prakash C. Sharma, Santosh Kumar Vishwakarma, Nirmal K. Gupta,
and Ashish Jain

Role of PID Control Techniques in Process Control System: A Review 659
Vandana Dubey, Harsh Goud, and Prakash C. Sharma

Investigating Cancer Survivability of Juvenile Lymphoma Cancer Patients Using Hard Voting Ensemble Technique 671
Amit Pandey, Rabira Galeta, and Tucha Kedir

Author Index 679

Editors and Contributors

About the Editors

Dr. Priyadarsi Nanda is Senior Lecturer at the University of Technology Sydney (UTS) with more than 27 years of experience specializing in research and development of cybersecurity, IoT security, Internet traffic engineering, wireless sensor network security, and many more related areas. His most significant work has been in the area of intrusion detection and prevention systems (IDS/IPS) using image processing techniques, Sybil attack detection in IoT-based applications, and intelligent firewall design. In cybersecurity research, he has published over 80 high-quality refereed research papers including transactions in computers, transactions in parallel processing and distributed systems (TPDS), future generations of computer systems (FGCS) as well as many ERA Tier A/A* conference articles. In 2017, his work in cybersecurity research has earned him and his team the prestigious Oman Research Council's National Award for best research. Dr. Nanda has successfully supervised 8 HDR at UTS (5 PhD + 3 Masters), and currently, supervising 8 Ph.D. students.

Vivek Kumar Verma is Assistant Professor (Senior Grade) in the Department of Information Technology, School of Computing & Information Technology, Manipal University Jaipur. His areas of expertise include data engineering, image processing, and natural language processing. He has published more than 35 research articles in peer-reviewed international journals, book series, and conference proceedings.

Sumit Srivastava is Professor of Information Technology with expertise in the domain of data analytics and image processing. He is also Senior Member of IEEE. He has published more than 100 research papers in peer-reviewed international journals, book series, and conference proceedings. He is having two edited volumes of Springer Nature Book series in his account. He has supervised more than 10 Ph.D. scholars in the areas of data science, machine learning, and natural language processing.

Rohit Kumar Gupta is Assistant Professor (Senior Grade) in the Department of Information Technology, School of Computer Science and IT, Manipal University Jaipur. He is currently pursuing a Ph.D. in the area of content distributed network from Malaviya National Institute of Technology, Jaipur. His research interests include cloud computing, network technologies, and IoT. He has published over 25 research articles in peer-reviewed international journals, book series, and conference proceedings.

Arka Prokash Mazumdar is Assistant Professor (Senior Grade) in the Department of Computer Science and Engineering, Malviya National Institute of Technology Jaipur, India. He has completed his Ph.D. from IIT Patna and M.Tech. from NIT Durgapur. His research interests include information-centric networks, cloud computing, and IoT. He has published over 50 research articles in peer-reviewed international journals, book series, and conference proceedings.

Contributors

Wadhah A. Abdulhussain Relational Machines Group, Human-Centred Computing Lab, School of Computing, Universiti Utara Malaysia, Sintok, Kedah, Malaysia

Ahsin Abid Department of CSE, Jahangirnagar University, Dhaka, Bangladesh

Md. Abu Ishak Mahy Department of CSE, Daffodil International University, Dhaka, Bangladesh

Garima Agarwal Manipal University Jaipur, Jaipur, India

Mahek Agarwal SCIT, Manipal University Jaipur, Jaipur, India

Pinky Agarwal Manipal University Jaipur, Jaipur, India

Aditya Aggarwal Netaji Subhas Institute of Technology, Delhi, India

Asif Ahmmed Department of CSE, Daffodil International University, Dhaka, Bangladesh

B. S. Aishwarya Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

R. Amulya Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Ho Nhat Anh Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam

Adarsh Ankur Department of Electronics and Communication Engineering, Amity School of Engineering and Technology, Amity University, Noida, Uttar Pradesh, India

Azizi Ab Aziz Relational Machines Group, Human-Centred Computing Lab, School of Computing, Universiti Utara Malaysia, Sintok, Kedah, Malaysia

Amit Kumar Bairwa Manipal University Jaipur, Jaipur-Ajmer Express Highway, Dehmi Kalan, Jaipur, Rajasthan, India

Praful Barekar Computer Technology Department, Yeshwantrao Chavan College of Engineering, Wanadongri, Nagpur, India

Joyanta Basak Department of CSE, Daffodil International University, Dhaka, Bangladesh

Joyanta Basu CDAC Kolkata, Salt Lake, Sector-V, Kolkata, India

Tapan Kumar Basu Department of Electrical Engineering, IIT, Kharagpur, West Bengal, India

Y. Bevish Jinila Sathyabama Institute of Science and Technology, Chennai, India

Darshana Bhagit SIES Graduate School of Technology, Navi Mumbai, India

Sumit Bhardwaj Department of Electronics and Communication Engineering, Amity School of Engineering and Technology, Amity University, Noida, Uttar Pradesh, India

Devershi Pallavi Bhatt University, Manipal University, Jaipur, Rajasthan, India

Al Amin Biswas Department of CSE, Jahangirnagar University, Dhaka, Bangladesh

Manoj K. Bohra School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Pinaki Chakraborty Netaji Subhas Institute of Technology, Delhi, India

Shampa Chakraverty Netaji Subhas Institute of Technology, Delhi, India

Tanishq Chamoli Chandigarh College Of Engineering And Technology (Degree Wing), Chandigarh, India

Sandeep Chaurasia Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur, India

Rajkishan Cherukuru SCIT, Manipal University Jaipur, Jaipur, India

A. Christy Sathyabama Institute of Science and Technology, Chennai, India

Parv Dahiya Amity University Haryana, Gurugram, Haryana, India

Lalit Damahe Computer Technology Department, Yeshwantrao Chavan College of Engineering, Wanadongri, Nagpur, India

Ramgopal Dhaka Department of Mathematics & Statistics, Manipal University Jaipur, Jaipur, Rajasthan, India

Leena Dihingia Indian Institute of Technology Guwahati, Guwahati, India

Saurabh Diwe Computer Technology Department, Yeshwantrao Chavan College of Engineering, Wanadongri, Nagpur, India

Adit Doshi Department of Computer Engineering, SCET, Surat, India

Vandana Dubey School of Computing & Information Technology, Manipal University Jaipur, Jaipur, India

Rabira Galeta College of Informatics, Bule Hora University, Bule Hora, Ethiopia

Sahil Garg Amity University Haryana, Gurugram, Haryana, India

Sonam Garg Chandigarh College Of Engineering And Technology (Degree Wing), Chandigarh, India

Aninda Ghosh Distronix, Kolkata, India

Harsh Goud Department of Electronics and Communication Engineering, Indian Institute of Information Technology Nagpur, Nagpur, India

Palash Goyal Mount Carmel School, Delhi, India

Somya Goyal Department of Computer and Communication Engineering, SCIT Manipal University Jaipur, Jaipur, Rajasthan, India

Ankit Gupta Chandigarh College Of Engineering And Technology (Degree Wing), Chandigarh, India

Gaurav Gupta Netaji Subhas Institute of Technology, Delhi, India

Nirmal K. Gupta School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Punit Gupta Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India

Rohit Kumar Gupta Manipal University Jaipur, Jaipur, Rajasthan, India

Shivani Gupta Vellore Institute of Technology, Chennai, TN, India

Iqbal Hasan Department of Computer Science, Jamia Millia Islamia University, New Delhi, India

Sk. Hasibul Islam Shad Department of CSE, Daffodil International University, Dhaka, Bangladesh

M. Imdadul Islam Department of CSE, Jahangirnagar University, Dhaka, Bangladesh

Anamika Jain Department of Mathematics & Statistics, Manipal University Jaipur, Jaipur, Rajasthan, India

Ashish Jain School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Kusum Lata Jain Manipal University Jaipur, Jaipur, RJ, India

Tarun Jain SCIT, Manipal University Jaipur, Jaipur, India

Divya Jangid Faculty of Mathematics and Computing, Department of Computer Science, Banasthali Vidyapith, Banasthali, Tonk, Rajasthan, India

Mahesh Jangid Department of CSE, Manipal University Jaipur, Jaipur, Rajasthan, India

Sandeep Joshi Manipal University Jaipur, Jaipur-Ajmer Express Highway, Dehmi Kalan, Jaipur, Rajasthan, India

Deborah T. Joy Amity University Gurgaon, Gurugram, Haryana, India

Md. Jueal Mia Department of CSE, Daffodil International University, Dhaka, Bangladesh

Sandeep Kakde Electronics Engineering Department, Yeshwantrao Chavan College of Engineering, Wanadongri, Nagpur, India

Shailesh Kamble Computer Technology Department, Yeshwantrao Chavan College of Engineering, Wanadongri, Nagpur, India

Kanika Netaji Subhas Institute of Technology, Delhi, India

Moksh Kant Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur, India

Muskan Kapadia Department of Computer Engineering, SCET, Surat, India

Rohan Kapoor Manipal University Jaipur, Jaipur, Rajasthan, India

Yash Kasliwal Department of Computer Science, Delhi Technological University, New Delhi, India

Rahul Katarya Department of Computer Science, Delhi Technological University, New Delhi, India

Mandeep Kaur Department of Computer Science and Engineering, Sharda University, Greater Noida, Uttar Pradesh, India

A. P. Kavya Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

B. Kavyashree Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Tucha Kedir College of Informatics, BuleHora University, Bule Hora, Ethiopia

Bui Thanh Khoa Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam

M. Kirthiga Sri Sairam Engineering College, Chennai, Tamil Nadu, India

Taikhum Kothambawala Department of Computer Science, Delhi Technological University, New Delhi, India

Sparsh Kotriwal School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Soummya Kulkarni SIES Graduate School of Technology, Navi Mumbai, India

Ashish Kumar SCIT, Manipal University Jaipur, Jaipur, India

Lalit Kumar Gazelle Information Technologies, Delhi, India

Rajesh Kumar College of Informatics, Bule Hora University, Bule Hora, Ethiopia

Rishav Kumar Mount Carmel School, Delhi, India

Saurabh Kumar Faculty of Physiotherapy, SGT University, Gurugram, Haryana, India

Shishir Kumar Department of Computer Science & Engineering, Jaypee University of Engineering and Technology, Guna, India

Sushil Kumar Jawaharlal Nehru University, New Delhi, India

Namrata Londhe JSPM'S Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Nguyen Minh Ly Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam

Sidhharth Mache Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, MH, India

Suhas Mache R. B. Attal Arts, Science and Commerce College Georai, Beed, MH, India

Manan Madan Netaji Subhas Institute of Technology, Delhi, India

Karina Maiyani JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Swanirbhar Majumder Department of Information Technology, Tripura University, Tripura, India

Ashish Malik Manipal University Jaipur, Jaipur, Rajasthan, India

Karan Malik Computer Science Engineering, USICT- Guru Gobind Singh Indraprastha University, Delhi, India

A. Malini Department of Computer Science, Thiagarajar College of Engineering, Madurai, India

D. T. Mane JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Deepak Mane JSPM'S Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Pratistha Mathur Manipal University Jaipur, Jaipur, India

Sourov Mazumder Department of CSE, Daffodil International University, Dhaka, Bangladesh

Shekhar Mishra Manipal University Jaipur, Jaipur, Rajasthan, India

Masooda Modak SIES Graduate School of Technology, Navi Mumbai, India

Abdella K. Mohammed Faculty of Informatics, IOT, Hawassa University, Hawassa, Ethiopia

Saurabh Mukherjee Faculty of Mathematics and Computing, Department of Computer Science, Banasthali Vidyapith, Banasthali, Tonk, Rajasthan, India

Ankit Mundra Manipal University Jaipur, Jaipur, Rajasthan, India

Kabir Nagpal Department of CSE, Manipal University Jaipur, Jaipur, Rajasthan, India

Nirali Nanavati Department of Computer Engineering, SCET, Surat, India

Parma Nand Department of Computer Science and Engineering, Sharda University, Greater Noida, Uttar Pradesh, India

Meenakshi Nawal Poornima College of Engineering, Rajasthan Technical University, Jaipur, RJ, India

S. Niharika Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Sunil Nimbhore Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, MH, India

K. L. Nirmal Raja Sri Sairam Engineering College, Chennai, Tamil Nadu, India

Bhoopendra Pachauri Department of Mathematics & Statistics, Manipal University Jaipur, Jaipur, Rajasthan, India

Amit Pandey College of Informatics, Bule Hora University, Bule Hora, Ethiopia

Hardik B. Pandit P.G. Department of Computer Science and Technology, Sardar Patel University, Anand, India

Parveen Guru Jambheshwar University of Science and Technology, Hisar, Haryana, India

Parveen Jawaharlal Nehru University, New Delhi, India

Dipjayaben Patel Vivekanand College For Advanced Computer And Information Science, Surat, Gujarat, India

Helly Patel Department of Computer Engineering, SCET, Surat, India

Mehul Patel JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Rikin Patel Department of Computer Engineering, SCET, Surat, India

Divya Patil JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Namita Patil JSPM'S Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Omkar Patil JSPM'S Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Tuhin Utsab Paul St. Xavier's University, Kolkata, India

Devi Prasad Sharma School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Soma Prathibha Department of Information Technology, Sri Sairam Engineering College, Chennai, Tamil Nadu, India

S. Prayla Shyry Sathyabama Institute of Science and Technology, Chennai, India

Jaya Krishna Raguru School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Nitin Rakesh Department of Computer Science and Engineering, Sharda University, Greater Noida, Uttar Pradesh, India

Priyanka Rishi Faculty of Physiotherapy, SGT University, Gurugram, Haryana, India

Nadia Afrin Ritu Department of CSE, Jahangirnagar University, Dhaka, Bangladesh

Sam Rizvi Department of Computer Science, Jamia Millia Islamia University, New Delhi, India

Sivapuram Sai Sanjith SCIT, Manipal University Jaipur, Jaipur, India

Priyankoo Sarmah Indian Institute of Technology Guwahati, Guwahati, India

Brijesh Satasiya Department of Computer Engineering, SCET, Surat, India

Abdus Sattar Daffodil International University, Dhaka, Shukrabad, Bangladesh

Madhavi Sawant JSPM's Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Siddhartha Saxena School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Priyanka Shah Kadi Sarva Vishwavidyalaya University, Gandhinagar, India

Venkatesh G. Shankar School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Ankush Sharma School of Physiotherapy, BUEST, Baddi, HP, India

Chitra Sharma Manipal University Jaipur, Jaipur, India

Harish Sharma Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur, India

Manoj K. Sharma School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Prakash C. Sharma School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Preeti Sharma University, Manipal University, Jaipur, Rajasthan, India

Utkarsh Sharma Department of Computer Science & Engineering, Jaypee University of Engineering and Technology, Guna, India

Rashmee Shrestha Department of Computer Science and Engineering, Sharda University, Greater Noida, Uttar Pradesh, India

Dhruv Shrivastav Mount Carmel School, Delhi, India

M. Shyamkumar Sri Sairam Engineering College, Chennai, Tamil Nadu, India

Lovedeep Singh Punjab Engineering College, Chandigarh, India

Neha Singh University School of Information, Communication and Technology, Guru Gobind Singh Indraprastha University, Dwarka, Delhi, India

Rishipal Singh Guru Jambheshwar University of Science and Technology, Hisar, Haryana, India

Shweta Sinha Amity University, Gurugram, Haryana, India

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

Salowa Binte Sohel Department of CSE, Daffodil International University, Dhaka, Bangladesh

P. V. Sridevi Andhra University College of Engineering (A), Andhra University, Visakhapatnam, AP, India

Utpal Srivastava Banasthali University, Jaipur, Rajasthan, India

Tipu Sultan Department of CSE, Daffodil International University, Dhaka, Bangladesh

Regidi Suneetha Andhra University College of Engineering (A), Andhra University, Visakhapatnam, AP, India

Md. Tariqul Islam Khulna University of Engineering and Technology (KUET), Khulna, Bangladesh

Vikas Thada Amity University Gurgaon, Gurugram, Haryana, India

Shreyas Thakur Department of Computer Science, Delhi Technological University, New Delhi, India

Pradeep Kumar Tiwari Manipal University Jaipur, Jaipur, India

Bhuvidha Singh Tomar Department of Electronics and Communication Engineering, Amity School of Engineering and Technology, Amity University, Noida, Uttar Pradesh, India

Nguyen Xuan Truong Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam

Abdur Nur Tusher Daffodil International University (DIU), Dhaka, Bangladesh

Sumaya Akter Usrika Daffodil International University, Dhaka, Shukrabad, Bangladesh

Mahima Manohar Varkhedi Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Shriya Verma Chandigarh College Of Engineering And Technology (Degree Wing), Chandigarh, India

Vivek Kumar Verma SCIT, Manipal University Jaipur, Jaipur, India;
Manipal University Jaipur, Jaipur, Rajasthan, India

Prashant Vidhate JSPM'S Rajarshi Shahu College of Engineering, Pune, Maharashtra, India

Deepali Virmani Department of Computer Science Engineering, Bhagwan Parshuram Institute of Technology, Rohini, Delhi, India

Santosh K. Vishwakarma School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Santosh Kumar Vishwakarma School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

Anju Yadav SCIT, Manipal University Jaipur, Jaipur, India

A. B. Yugakiruthika Department of Computer Science, Thiagarajar College of Engineering, Madurai, India

Using Machine Learning, Image Processing and Neural Networks to Sense Bullying in K-12 Schools: Enhanced

Lalit Kumar, Palash Goyal, Karan Malik, Rishav Kumar, and Dhruv Shrivastav

Abstract We all have heard about bullying, and we know that it is an immense challenge that schools have to tackle. Many lives have been ruined due to bullying, and the fear it implants into students' mind has caused many of them to go into depression which can lead to suicide. Traditional methods (National Academies of Sciences, Engineering, and Medicine, in Preventing bullying through science, policy, and practice. The National Academies Press, Washington, DC, [1]) need to be accompanied with modern technology to make the method more effective and efficient. If real-time alerts are sent to school staff, they can identify the perpetuator and extricate the victim swiftly. In this proposed method, an AI-based solution is implemented to monitor students using standard school surveillance technologies and CCTV to maintain a decorum and safe environment in the school premise. Also the proposed method utilizes other unstructured sources such as attendance records, social media activity and general nature of the students to deliver quick response. Artificial intelligence (AI) techniques like convolutional neural networks (CNNs), which include image-processing capabilities, logistic regression methods, long short-term memory (LSTM) and pre-trained model Darknet-19 are used for classification. Further, the model also included sentiment analysis to identify commonly used abuse terms and noisy labels to improve overall model accuracy. The model has been trained and validated with the realistic data from all the sources mentioned and has achieved the classification accuracy of 87% for detecting any sign of bullying.

Keywords AI · CNN · Class entropy loss · Data-pre-processing · Data pipeline · Facial recognition · NLP · NN · Sigmoid · Sentiment analysis · LSTM · Darknet-19

L. Kumar
Gazelle Information Technologies, Delhi, India

P. Goyal (✉) · R. Kumar · D. Shrivastav
Mount Carmel School, Delhi, India
e-mail: palashgoyal1608@mountcarmeldelhi.com

K. Malik
Computer Science Engineering, USICT- Guru Gobind Singh Indraprastha University, Delhi, India

P. Nanda et al. (eds.), *Data Engineering for Smart Systems*, Lecture Notes in Networks and Systems 238, https://doi.org/10.1007/978-981-16-2641-8_1

1 Introduction

What is bullying? For starters, bullying is aggressive behavior among teenagers. This comes from many factors which involve ego, power imbalance, upbringing, etc. This behavior has no bounds and can impact a victim's physical and mental health conditions. Bullying can take many forms from verbal to physical. With the introduction of social media, social bullying, a new type of bullying which involves harming one's reputation or relationships, has sprung up. A study was conducted by Symantec Reports with the help of parents of many victims. They noted that almost 24% of the students were involved in some shape or form of bullying [2]. Despite such a high percentage of victims, most methods to keep bullying in check have not borne any fruit and mitigation against bullying remains an enigma.

The traditional methods with their limited use of technology are proved not be very effective. It is time to integrate modern technology to develop more intelligent solutions. The proposed method in this paper will use specific elements of AI tools which exploit the already-available school infrastructure to make it a new means of keeping taps on bullying.

2 Methodology of Research and Solution Development

First of all, we have identified all the physical, mental, emotional, social and cyberbullying types and parameters that are prevalent. For this, extensive study of existing literature has been done [1–3]. Subsequently, these parameters were studied to identify how they can be objectively analyzed. This involved identifying the data sources that can give input signals for that bullying parameter. Once the data sources were identified, actual data was recovered from these sources. Again, existing literature was studied [4–7] extensively to identify AI/ML algorithms that can be used to classify them with binary outputs linked to whether this is a bullying situation or not. In case bullying is detected, our system detects and recognizes faces of those involved and sends real-time alerts to the authorities, with the identity of the perpetrator [8, 9]. The authors then put together all these elements to develop a working model of their solution with a proof of concept done in a school. The research methodology followed in this is as given below in Fig. 1.

2.1 Bullying Parameters/Features

To identify the signs of bullying, we have taken multiple parameters under consideration through which we can filter-out real bullies from other behaviors which could be mistaken for bullying. Features like rude behavior, troubling authorities regularly, low attendance, use of foul language, low scores, drug use, explicit content, etc.

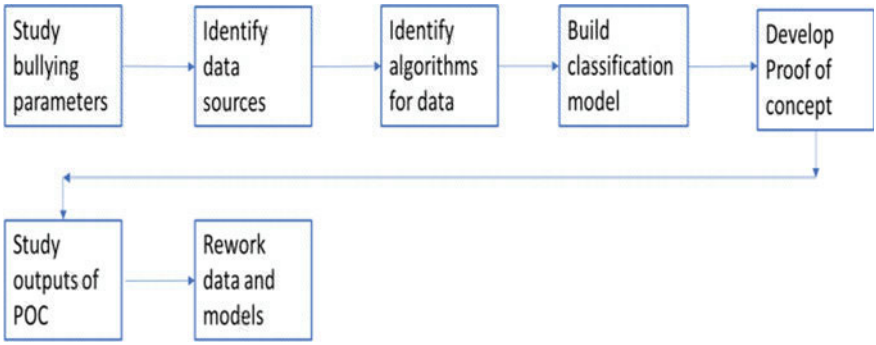


Fig. 1 Research methodology

Table 1 How to identify a victim of bullying

Victim parameter	Data source
Show deteriorating grades	Academic report
Regular absenteeism	Attendance records
Being picked on, pushed, punched, etc. (physically harassed)	CCTV images
Downfall of social skills	Registration for school events and extracurricular groups
Suffering from learning or mental disabilities	Counselor reports
Unwilling to go to school regularly	Attendance records
Random bruises, missing belongings or torn clothing	CCTV images
Prone to attacks of anxiety	School medical reports
Alone at lunch-breaks	CCTV images
Experiences regular nightmares	Medical room report
Starts bullying younger or weaker kids to vent out the frustration	CCTV images

are used to filter the data. In the proposed model, it is planned to integrate AI tools with school infrastructure and data like cameras with microphones, student portals containing community forums where students and teachers interact, attendance of students, their scorecards, etc. Data used for analysis and a complete infrastructural setup is discussed below in Tables 1 and 2.

Table 2 Bullying perpetrator parameters and data sources

Perpetrator parameters	Data source
Rude behavior with students as well as teachers	CCTV images, audio, school report
Low grades	School report
Lacks empathy or guilt	School report
Feeling of entitlement because of being good in school, sports or belonging to a prominent family	School report
Short-tempered and having emotional outbursts	CCTV images, audio, school report
Usually popular or among a big group	CCTV images, audio, school report
Regularly get into trouble with authority	School report

2.2 Integration with School Infrastructure

Recordings of playground, common-area (like corridor, locker rooms, etc.) and classes will be taken along with their audios via cameras that are installed. More detailed information associated with students will be extracted from student portals. This will further help in understanding and analyzing student behavior and personality. This raw information is then reworked into structured data, which will supplement the learning algorithm in predictions and analysis. Based on this, appropriate action can be taken by school authorities. This method can also be inverted and be used to discern the victims of harassment.

Data Analysis:

Photos uploaded by the students on community forums along with acquired video footage from CCTV are split into images (see Figs. 2 and 3) and ran through the algorithm to identify: drugs, number of faces, anxiety attacks, crying, isolation, fighting, torn clothes, bruises, sleeping, smoking, hard drinks, gore, explicit and adult content. The image classifier uses CNN and pre-trained network called Darknet-19 (trained

Fig. 2 Sample images (in case of bullying) from dataset

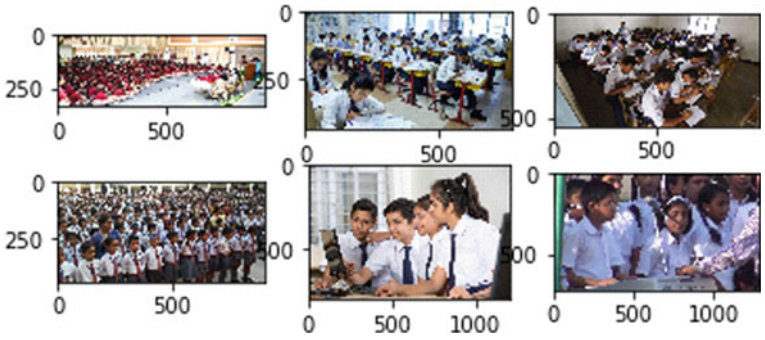


Fig. 3 Sample images (non-bullying) from dataset

on ImageNet dataset), along with LSTM which provides the capability to process sequence of image data with feedback connections. When a bullying scenario is detected by the CNN-LSTM architecture, faces of all those involved or present in the particular frame are detected using the popular object detection algorithm You Only Look Once (YOLO). These faces are input into another simple CNN, namely a Siamese network, which tries to identify the perpetrator by matching the input with the student database already present.

Audios which are mapped to text using Google Cloud Speech API alongside comments at student community are further used to determine the following features: tone, amplitude and pitch of voice, language used (explicit or not), uppercase text, text length and sentiments, threatening statements, trolling, unpleasant comments and distasteful words. Other attributes like low attendance, incompetent grades, enrolment in extracurricular activities, teacher–student interaction, frequency of councilor appointments and behavior report by staff will also be considered. Data for these inputs is converted from physical form into digital form first.

2.3 Techniques Involved

Convolution neural network is a type of neural network (NN) that provides the capability to convert pixels into well-structured data [5]. CNNs replicate the function of the frontal lobe of the human brain (cerebral cortex), which is responsible for processing audio visual stimulus in humans. To process images, CCTV footage is spliced into still images and then each frame is analyzed to extract the crucial and important features which can be further analyzed for more refined results.

Proposed network architecture for video classification is shown in Fig. 4. It has been already shown that adding LSTM (which extract global temporal features) after CNN, the local temporal features obtained from optical flow are also of great importance [10].

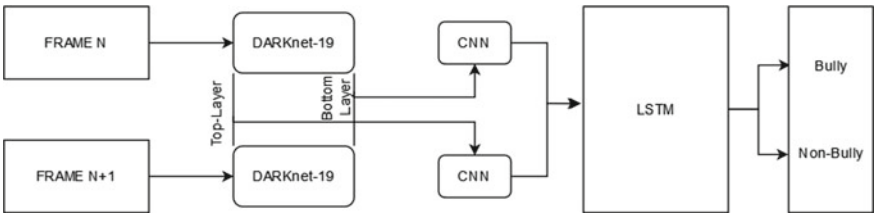


Fig. 4 Architecture of model

Optical flow is due to displacement of boundaries or movement of object across two consecutive frames. Optical flow is best used for action recognition, so its functionality is mimicked by taking two consecutive frames as input to training model. These consecutive frames are fetched to pre-trained CNN model (Darknet-19). Then, both of the frame values from the bottom layer of pre-trained model are fed as an input to additional CNN. Additional CNN network is now to supposed to learn from the local motion features including the invariant features by comparing the both frame. The top layer of the pre-trained network is also fetched to another additional CNN to learn from the comparison of high-level features of both the frames. Furthermore, the output of both the additional CNN is fully connected with LSTM cell, which further classifies the images as bully and non-bully.

When the image is classified as bully, the YOLO algorithm, which is an object detection algorithm, is used to detect all the faces present in the "bully" frame, using a vector of bounding boxes (see Fig. 5). It uses a convolutional neural network (CNN) to complete this task. The structure of the CNN used and a brief working is shown in Fig. 5. The input image is divided into grids or regions, and bounding boxes are predicted for each of these regions. The prediction vectors for each region contain six elements, namely [x, y, w, h, c, p], where x, y, w and h are the coordinates of the bounding boxes, c is the confidence and p is the probability of occurrence of faces, i.e., if the algorithm fails to detect any faces, p becomes equal to 0 and the remaining values are not defined.

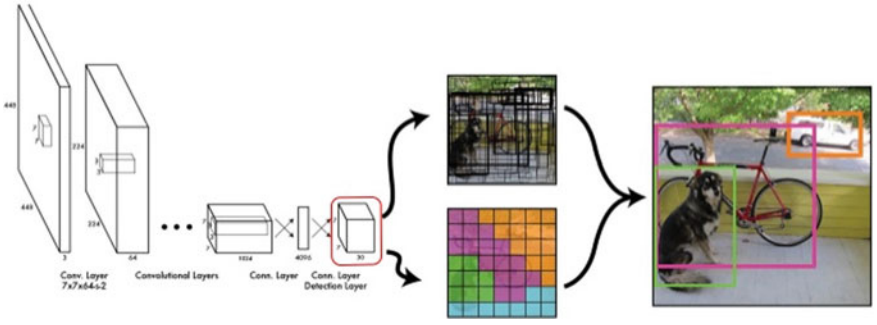


Fig. 5 YOLO algorithm

For unambiguous and noise-free detection, threshold values are set, in order to filter-out disturbances. Only regions that have their probability (p) value exceeding the threshold are considered. To further improve quality of predictions, Intersection over Union (IoU) and non-max suppression are performed. Non-max suppression helps in eliminating any bounding regions that might have detected an already detected feature or face. The final bounding boxes received after the complete procedure are sent as inputs into a Siamese network for facial recognition.

Since only one picture of each student will be present in the database, training a deep learning algorithm to map the received images to one stored in the dataset will not be feasible. Due to this, a Siamese network, another variation of the convolutional neural network that is a one-shot learning algorithm is used. It gives high performance in applications and scenarios where there is not a large amount of data belonging to each data class.

A Siamese network consists of two parallel neural networks, with the same weights and parameters. It consists of convolutional, pooling and dense layers just like a regular CNN, but what distinguishes it from it is that it is bereft of a softmax/sigmoid activation at the last layer. Thus, the encoding of an image is received as an output instead of a category. The contrastive loss or difference between encodings received from the two parallel networks is calculated to ascertain the similarity between the input images.

The faces inside the bounding boxes predicted by the YOLO algorithm are extracted and provided as input to one of the two parallel networks. The other network receives images from the database of students, and the difference of the received encodings in calculated. The image having the least difference is considered to be "identified" as one of those present in the bullying frame. To fasten the process and reduce computational needs, the encodings of all students in the database were generated in advance using the same Siamese network, and only difference calculation was done during the test phase. Immediate alerts, including the identity of those involved, are then sent to the authorities. Figure 6 displays the working of a Siamese network for facial recognition.

For audio analysis, we have mapped CCTV voice output with Google Cloud Speech API which uses CNN and provides real-time streaming of speech recognition and conversion from audio to text. This lies within our CCTV and Google Cloud Storage for sentiment analysis to be done on it.

For text analysis, mentioned words are mapped with sentiment analysis dictionary ("Liu and Hu opinion lexicon" containing 6800 positive and negative words [4]). To extract feature from the text, multinomial Naïve Bayes is used (refer Fig. 7). It helps in classification of bad words and rude comments from audio-to-text converted files as well as from student's community portal including explicit and vulgar remarks, text length as well as usage of upper-case letters in community forums, etc. All such attributes are then listed for feature extraction by the model.

Even data like medical records, attendance, grades and teachers remark from student portals is used and processed for feature extraction. Logistic regression modeling is used for this analysis.