Proceedings Of

International Conference on Advances in Computing and Communication (ICACC-2021)

May 19-20, 2021

Organized by

Department of Information Technology

Dr. AKHILESH DAS GUPTA INSTITUTE OF TECHNOLOGY & MANAGEMENT, NEW DELHI

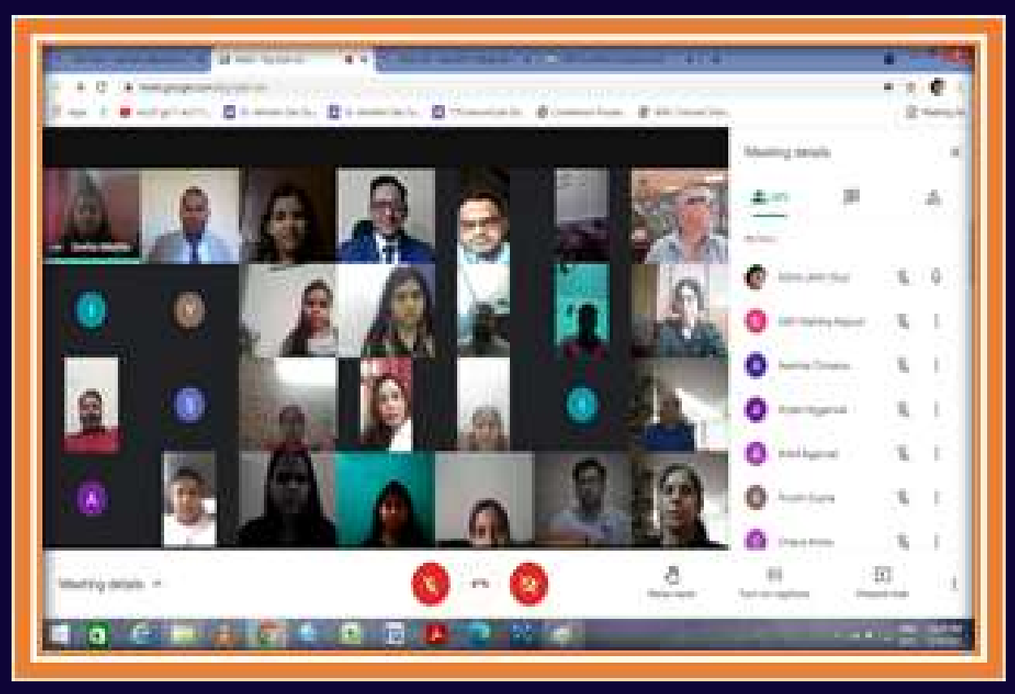
(Formerly Northern India Engineering College)

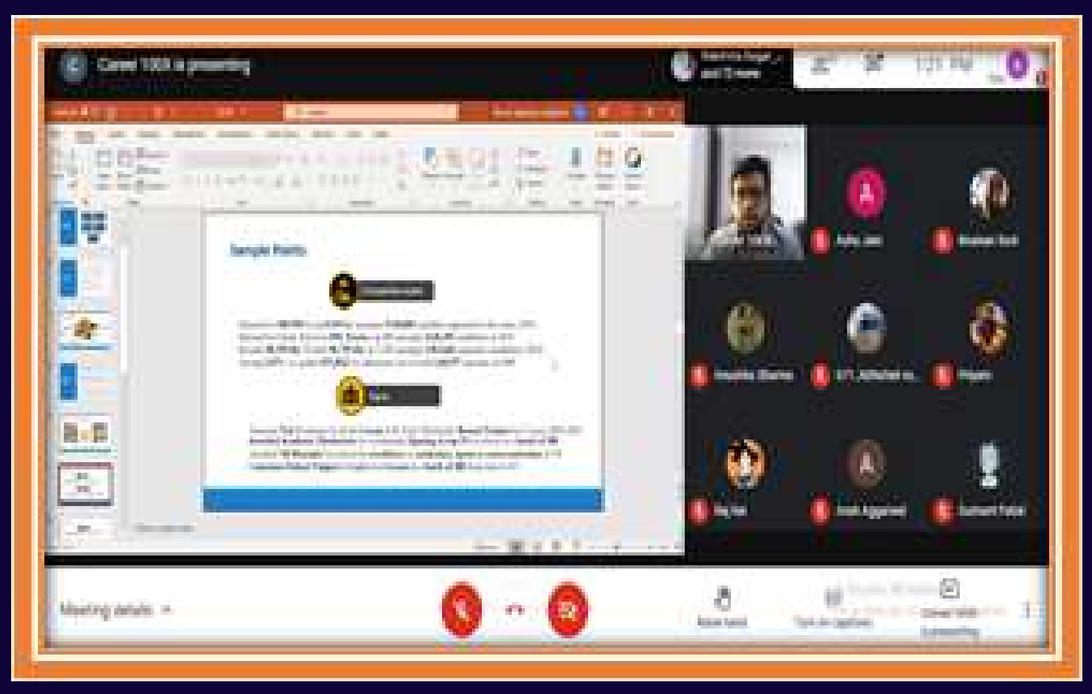
(Babu Banarasi Das Group of Educational Institutions)

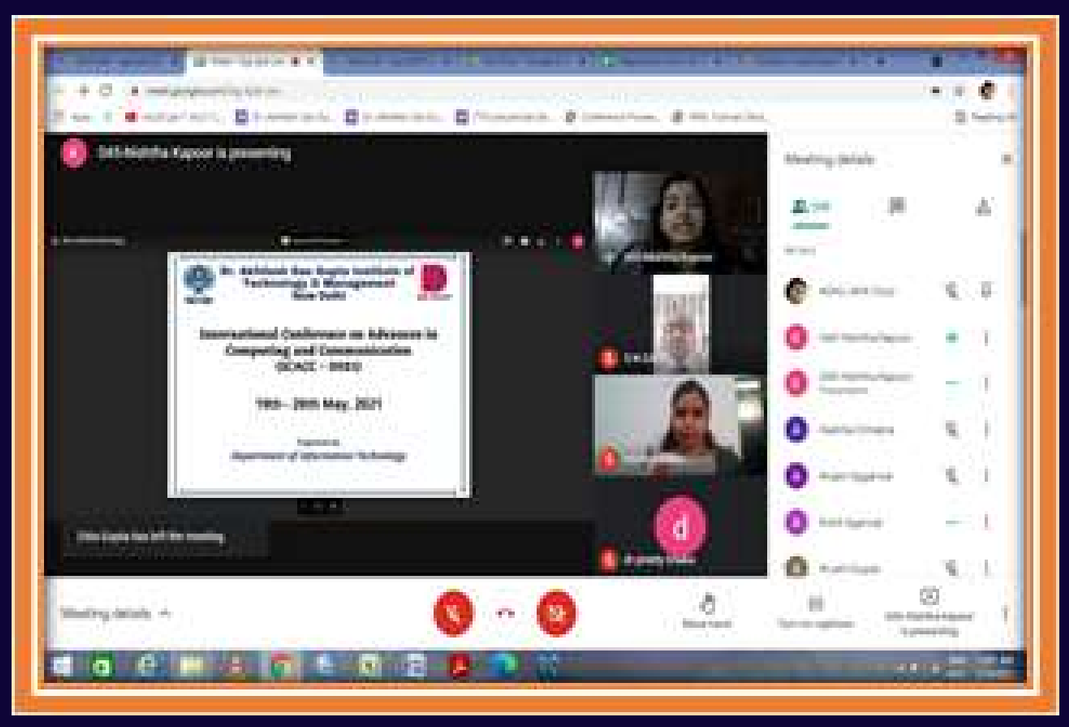
Affiliated to Guru Gobind Singh Indraprastha University, New Delhi

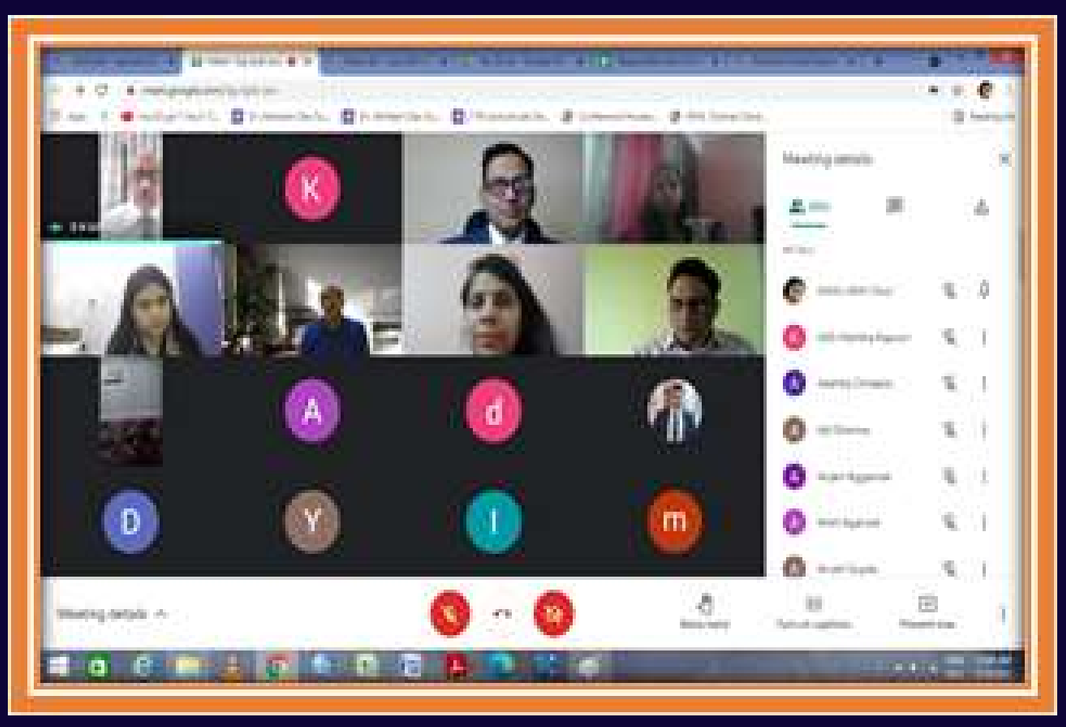
&

Approved by AICTE, New Delhi






International Conference on Advances in Computing and Communication
19th - 20th May, 2021



INTERNATIONAL CONFERENCE

ON

ADVANCES IN COMPUTING &

COMMUNICATION

(ICACC-2021)

19th - 20th May, 2021

Organized By

Department of Information Technology

Dr. Akhilesh Das Gupta Institute of Technology & Management

(Unit of Babu Banarasi Das Educational Group)

Approved by AICTE, Accredited by NAAC, Affiliated to GGSIP University, New Delhi

FC-26, Shastri Park, New Delhi-110053

Ph: 01149905900-99

Website: www.adgitmdelhi.ac.in

Declaration

All data, views, opinion etc. being published in this proceeding are the sole responsibility of the Authors. Neither the publisher nor the organizer of the conference will be responsible in any way for them.

Title: International Conference on Advances in Computing and Communication (ICACC – 2021)

Editor's Name: Prof. (Dr.) Prashant Singh, Ms. Charul Dewan, Ms. Ashu Jain

Published by: Department of Information Technology

Publisher's Address: Dr. Akhilesh Das Gupta Institute of Technology & Management, FC-26, Shastri Park, New Delhi-110053

Printer's Details: Conference Proceedings

Edition Details: I

ISBN: 978-93-5493-753-8

Organizing Committees

CHIEF-PATRONS

Mrs. Alka Das Gupta, Chairperson, BBD Group

Mr. Viraj Sagar Das, President, BBD Group

Ms. Sonakshi Das, Vice President, BBD Group

PATRONS

Shri S.N.Garg, C.E.O
Prof. (Dr.) Sanjay Kumar, Director
Prof. (Dr.) Yamini S., Principal & IQAC Director
Ms. Pankhuri Aggarwal, Director, HR

CONVENER

Prof. (Dr.) Prashant Singh

Dean (Academics) & Head (IT)

CO-CONVENER

Ms. Charul Dewan

Ms. Ashu Jain

ADVISORY COMMITTEE

- Prof. David Brown, Nottingham Trent University, United Kingdom
- Dr. Omprakash Kaiwartya, Nottingham Trent University, United Kingdom
- Dr. Mukesh Prasad, UTS, Australia Dr. Abbas Mehrabidavoodabadi, Northumbria University, United Kingdom
- Dr Mufti Mahmud, Nottingham Trent University, United Kingdom
- Dr. Ahmad Mohammad Saleh Khasawneh, Amman Arab University (AAU), Amman, Jordan Prof. Yue Cao, Wuhan University, China
- Dr. Rupak Kharel, Manchester Metropolitan University, United Kingdom
- Dr. Geetika Aggarwal, Nottingham Trent University, United Kingdom
- Prof. D. K. Lobiyal, JNU, Delhi
- Dr. T. Ganesh Kumar, Galgotias University, UP
- Prof. Devendra Tayal, IGDTUW, Delhi
- Prof. S R N Reddy, IGDTUW, Delhi
- Dr. Subarahmanyam, IIT, Ropar
- Dr. R Bala Subramanyam, IIT, Roorkee
- Dr. Deepak Kumar Sharma, NSUT, Delhi
- Dr. Nanhay Singh, NSUT(East Campus), Delhi
- Dr. Sanjay Kumar Dhunrandher, NSUT, Delhi
- Dr. Sushil Kumar, SC & SS, JNU, Delhi
- Dr. Shailender Kumar, DTU, Delhi
- Dr. K. Sampath Kumar, Galgotias University, UP

ADGITM ADVISORY COMMITTEE

Dr. Pratul Arvind
Dr. Anupam Sharma
Dr. Rajeev Sharma
Dr. Deepak Bhardwaj
Dr. Mohit Singh
Mr. Ajeet Kumar
Ms. Maninder Kaur
Dr. Kavita Gupta
Dr. Saurabh Gupta
Dr. Surender Dhiman
Dr. Divya Gangwar
Dr. Preety Verma Dhaka
Mr. Neeraj Kumar

ORGANIZING COMMITTEE

Mr. Ankit Agarwal
Mr. Jogender Kaushik
Mr. Amit Pandey
Ms. Monica
Ms. Aashita Chhabra
Ms. Shimpy Goyal
Ms. Anjani Gupta
Ms. Saijal Gupta
Mr. Devender Banga
Ms. Kajal Kaul
Ms. Sneha Mishra
Ms. Shuchi Srivastava
Mr. Dhyanendra Jain
Mr. Ujjval Jain
Ms. Lalita Luthra
Ms. Gunjan Chugh
Ms. Neha Sharma
Ms. Princy Jain
Mr. Gourav Sharma
Ms. Arushi Gupta
Ms. Priyanka Singh
Ms. Deepika Tyagi
Mr. Mohit Dayal

Editorial Board

With immense pleasure and delight, we would like to announce a two-day International Conference on Advances in Computing and Communications (ICACC-2021) conducted by Department of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology & Management, New Delhi on 19th -20th May, 2021.

The conference aims to provide a platform for students, research scholars, academicians, and industrialists to meet and share their views on the recent technologies for the overall growth in Computing and Communications technologies. The purpose is also to promote the research and practice of new strategies, tools, techniques and technologies for the design, development and implementation of Information system.

Computing and Communications technologies are progressively evolving by covering the major areas like artificial intelligence, embedded systems, mobile and wireless communications, image processing, grid and cloud computing, semantic web, user experience and security and so on.

We believe that this conference would highlight many new frontiers of knowledge and provide ideas of research.

We are grateful to the management and various faculty members of Dr. Akhilesh Das Gupta Institute of Technology & Management, New Delhi and to those who have directly or indirectly helped us in the successful conduction of this conference.

We look forward to suggestions and recommendations from all sections of researchers, academicians, students and delegates to make this endeavor fruitful for all.

Ms. Charul Dewan
Co-Convener
ICACC-2021

Ms. Ashu Jain
Co-Convener
ICACC-2021

Prof. (Dr.) Prashant Singh
Convener
ICACC-2021

Prof. (Dr.) Sanjay Kumar
Patron
ICACC-2021

Hon'ble Chairperson, BBD Group

Mrs. Alka Das Gupta

L.L.B, MBA

Message

It is a matter of great pleasure that our institute–Dr. Akhilesh Das Gupta Institute of Technology and Management New Delhi, is organizing *an International Conference on Advances in Computing and Communications* on 19th & 20th May,2021 and thus providing a platform for researchers and scholars to share pioneering ideas and technologies.

Computing and Communications offers transformative, integrated, and robust information technology services, systems and infrastructures in support of trifold mission of teaching, research, and public service.

I believe that this International conference would offer an excellent opportunity to the participants to deliberate on this significant theme which in turn would help in extending the opportunities in the field of *Digital India* offered by *Advancement in Computing and Communication.*

I extend my best wishes and hearty congratulations to the organizing team.

Mrs. Alka Das Gupta
Hon'ble Chairperson
BBD Group

Hon'ble Vice-Chairman, BBD Group

Mr. Viraj Sagar Das

Message

Dr. Akhilesh Das Gupta Institute of Technology and Management, New Delhi has a rich tradition of achieving academic excellence and value-based education by providing a conducive environment to its students, staff and faculty members for their continuous growth and development. This is a matter of great pride that our institute's academic and research environment is advancing with each passing day.

With this aim in mind the Institute is conducting an *International conference on Advances in Computing and Communications* on 19th & 20th May, 2021.

Computing and Communications provide technology-based services and support with a commitment to diversity within an environment that fosters collaboration and collegiality.

I hope that this deliberated international conference on such an important theme will inculcate latest technical knowledge and competencies by extending opportunities that benefit everyone.

I extend my heartiest congratulations to the entire organizing team for making the event a resounding success.

Mr. Viraj Sagar Das
Hon'ble Vice-Chairman
BBD Group

Chief Executive Officer, ADGITM

Sh. S. N. Garg

Message

Information Technology is the need of the day in today's competitive world. Computing & communication plays a major role in Information Technology & Communication sector.

It is an enlightening and scrupulous endeavour of Department of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology and Management, Delhi to organize an International conference on such an enticing topic and providing an opportunity to the participants and delegates to upgrade their knowledge. This *International conference on Advances in Computing and Communications* on 19th & 20th May, 2021 brings into light all major changes taken place in latest technologies like Big Data, Internet of Things (IoT), Grid and Cloud computing, Image Processing, Software Engineering, Mobile and Wireless communications, Web semantic and so on.

I hope such conferences will go a long way to benefit not only participants but others too connected to it.

We thank Hon'ble Chairperson and Hon'ble President of BBD Group Society for their immense support. I would also like to appreciate the active participation of faculty and students for their efforts to make this International conference a grand success.

Sh. S.N. Garg
CEO
ADGITM

Director, ADGITM

Prof. (Dr.) Sanjay Kumar
Ph.D, M.Tech, B.E.

Message

Dr. Akhilesh Das Gupta Institute of Technology and Management (ADGITM), New Delhi is a thriving Institute that encourages a four-way partnership among Students, Parents, Institute and Industry. It gives me immense pleasure to announce that ADGITM is organizing an *International conference on Advances in Computing and Communications* on 19th & 20th May, 2021. The conference is an attempt to provide a wholesome experience and exposure to all participants from academia and industries with an assurance of their contributions in nation building and World progress.

The objective of the conference is to bring eminent students, research scholars, academicians, industrialists and experts from different technical domains to update theirknowledge and explore new horizons in the field of Computing and Communications.

I am sure that the contributions in the form of research papers will enrich the knowledge of all participating in the conference and will benefit the society as a whole.

I congratulate the conference organizing team for their hard work and best of luck for their future endeavors.

Prof. (Dr.) Sanjay Kumar
Director
ADGITM

Dean (Academics) & Head (Information Technology)

Prof. (Dr.) Prashant Singh
Ph.D, M.Tech, B.E.

Message

Teaching & learning agility is most essential in order to keep up with the speed of change across the competitive world by knowing latest trends, techniques and technologies for the design, development and implementation of Information systems. It is through the spirit of teamwork and strong sense of vision that we stride towards our mission of boosting academic as well as research capabilities by organizing an *International conference on Advances in Computing and Communications* on 19th& 20th May 2021. We pave the way towards creating value for all our stakeholders and truly honor our vision towards the betterment of society.

I am grateful to the chief patrons - Hon'ble Mrs. Alka Das Gupta, Chairperson, BBD Group, and Hon'ble Mr. Viraj Sagar Das, President, BBD Group, for valuable guidance, financial support and cooperation extended to us for the successful conduction of this International Conference.

I would like to appreciate the organizing committee for their marvelous effort in successfully organizing the conference. I shall also like to convey my deep appreciation to all participants for their applause worthy efforts and quality papers.

Dr. Prashant Singh
Convener
ICACC – 2021

Contents

An insight on Computer Vision Deploying Image Processing and Object Detection: Applications and Future Research Opportunities

Geetika Aggarwal[1], Neil Mansfield[2] ,Frederique Vanheusden[3]
Department of Engineering, School of Science and Technology
Nottingham Trent University, Clifton Campus
Nottingham, UK, NG11 8NS
[1]geetika.aggarwal@ntu.ac.uk;
[2]neil.mansfield@ntu.ac.uk

***Abstract*— The Computer vision (CV) is a study to understand the functionality of computers to collect information of images or videos captured by a camera while the processing of images or frames to extract the original data is known as image processing. In the object detection, computer vision is extensively used to depict the distance between objects, training, motion of objects through an image or video. The information obtained from computer vision can be used for image processing, object detection, tracking etc. Over the years with the technology advancement and inbuilt cameras in smart devices, laptops, iPads,etc., computer vision has found its existence in numerous areas. This paper will provide an insight on computer vision deploying image processing and object detection. The potential applications of computer vision are also discussed in this research paper along with challenges and future research opportunities.**

***Keywords*—Computer Vision, Image Processing, Object Detection**

I. INTRODUCTION

In Computer Vision (CV), algorithms are developed to instruct computers to perform tasks through the knowledge of the content's specific images, frames, and videos. Computer vision is a subset of both artificial intelligence (AI) and Machine learning (ML), shown in figure 1. The algorithms of both AI and ML combined with computer vision such as object detection and image processing are deployed to recreate human vision. Nowadays the increased use of mobile cameras means a steady stream of images and videos, and the technology of Computer Vision has become readily available, making it even more attractive to companies, organizations and institutions [1]. In the 1950s, when early neural networks started to detect the edges of objects and to organize them by their types, Computer Vision took its first steps. The first commercial Computer Vision systems were used in the 1970s, using optical character recognition (OCR) to read written text for the blind. Huge collections of images became available online for review as the internet evolved in the 1990s, driving the development of facial recognition programmes. Computer vision is an interdisciplinary scientific field that deals with how visual images or videos can help computers achieve high- level understanding [2-3]. In this context, comprehension implies the transformation of visual images (the retina input) into world explanations that make sense of thought processes and can evoke effective action. This interpretation of images can be seen as the disengagement of symbolic knowledge from image data using models developed with the assistance of geometry, physics, statistics, and the theory of learning [4- 6].

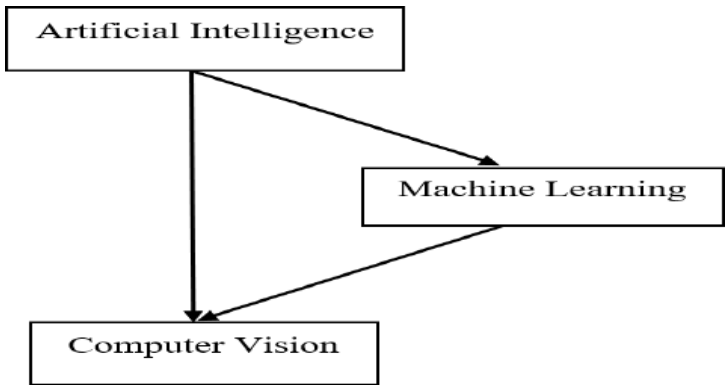


Fig.1 Family tree of computer vision

To construct computer vision systems, the technical discipline of computer vision aims to apply its theories and models. Site reconstruction, event identification, video tracking, object recognition, 3D pose estimation, learning, indexing, motion estimation, visual servo, 3D scene modelling, and image recovery are sub-domains of computer vision [7]. Fast image acquisition enables 3D measurement and feature tracking to be realized when paired with a high-speed projector [8] Egocentric vision systems are composed of a wearable camera that takes photographs from a first-person perspective automatically.

Over the past few decades, the researchers have been working on to find the directions to educate computers or machines to effectively communicate deploying visual data. In 1959, the first experimentation was carried by neurophysiologists using cat and set of images, aiming to correlate the response of the brain of cat. It was discovered that the brain of the cat responded first to sharp lines or edges in an image, resulting to conclude that processing of images

corresponds to edges, lines, and shapes [9]. Thereafter, following the interpretation on the initial concept of image processing digitization of images was possible due to development of image scanning technology. With the substantial research into image processing and computer vision, by 1963, the 2D (two-dimensional) images could be translated to 3D (three-dimensional) images. From 1960s, until 1970s, AI deploying computer vision has been extensively used in an academics to resolve the problems related to human vision. Furthermore, from 1974 recognition of printed text and translating the hand-written text through optical character recognition (OCR) technology and intelligent character recognition (ICR) respectively are effectively performed in numerous applications. In 1982, neuroscientist David Marr established that computer vision works hierarchically and introduced algorithms for machines to detect edges, corners, curves and similar basic shapes. Concurrently, computer scientist Kunihiko Fukushima developed a network of cells that could recognize patterns and object detection [11,12].

This paper is divided into five different sections. Section 2 gives an overview of computer vision, image processing and object detection. The applications of computer vision are listed in section 3, where they are critically discussed and compared. The paper concludes its review in section 4. The challenges and future research opportunities for computer vision are mentioned and discussed in section 5.

II OVERVIEW

The use of computer vision to analyze the data from images and videos deploying cameras has been expanded to institutions, organizations, and companies.

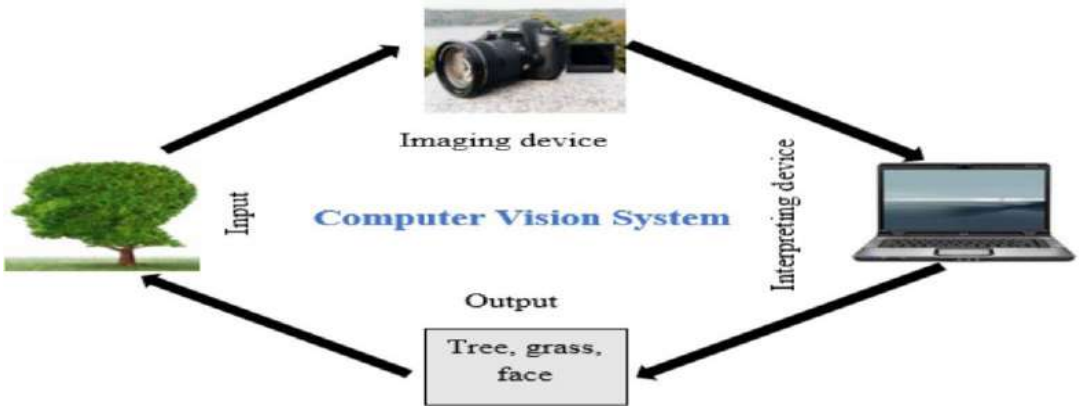


Fig.2 Computer Vision system.

The simplest example of computer vision system is shown in figure 2, where the input of the image is captured by a camera, analyzed through an interpreting device, and then producing the output [13]. The input to the camera could be a video for real-time application or an image, for the computers to recreate using computer vision. The embedded systems such as micron troller, ARM processor, hardware design could be an interpreting device too, which are able to process the algorithms and extract the information from image sand videos. In figure 2, camera capture an image and when interpreted the output is, tree, grass, and face, thus resulting in effective working system model of computer vision. In computer vision, an image in 2D is made of a combination of rows and columns [13,14] shown in figure 3. The intersection of row and column in an image, resulting in formation of small blocks in an image are known as pixels.

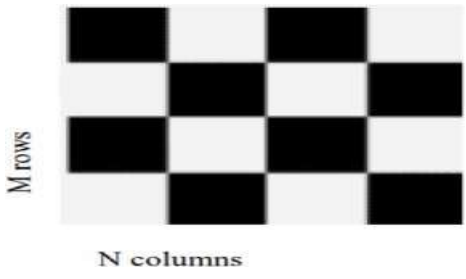


Fig. 3 Image in the form a matrix of rows and columns.

Pixels in an image is basic or smallest unit of an image, combined to form an image. The pixel size in an image is crucial as depending on the size of the pixel in an image the formation of rows and columns in an image occurs within a specified height and width of the respective image [14,15]. For example, the smaller the size of the pixel, more rows, and columns in a specific image, larger the pixel size, a smaller number of rows and columns. The application of pixel size if beneficial in data rate and communication systems, because with the smaller size of pixel data rate is enhanced in comparison to larger pixel size in an image [16-18]. In figure 4, a data stream sent through communication converted to pixels of black and white are shown. Black pixels are represented by '0' and white pixels are represented by '1' respectively.

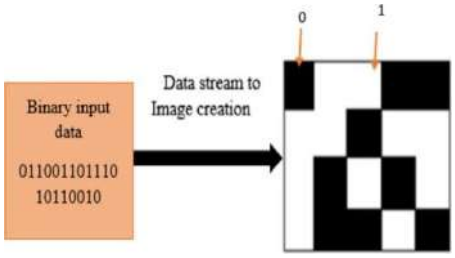


Fig. 4 Binary data to Image creation.

It is clear from figure 4, that binary stream when communicated using images is converted to image results in a matrix of rows and columns forming a 2D array or matrix with black pixels representing '0' and white pixels representing '1' within the two-color bits i.e., 0 and 1. Theis formation of matrix could be used to communicate wirelessly using QR code in object detection, motion tracking, etc. Computer vision deploying images and object detection though the knowledge of images and videos is deployed in several application areas discussed in next section.

III COMPUTER VISION APPLICATIONS

Due to the advancement in technology such as inbuilt cameras in smart phones, laptops, smart watches etc, the research on computer vision, image processing and object detection has gained substantial interest. The quest of understanding and study of images captured has expanded the expanded the computer vison to different domains. automates processes in a way that not only reduces human effort but also provides us with solutions to the task that could never have been solved by the limitations of the human vision. Some potential applications of computer vision, are mentioned below:

A. Healthcare

Many diagnoses in healthcare rely on images, hence, there exists a great potential of computer vision in healthcare in saving lives and aiding in treatment. Often, the medical images are difficult to categorize, therefore the deployment of computer vison algorithms along with artificial intelligence and machine learning can help medical professionals by enhancing images, and also by automatic screening such that more time is available for scrutinizing possible positive tests [21]. Computer vision in healthcare can be used for surgical simulation and surgical assistance. The technology can help surgeons with complicated decisions, especially during laparoscopic surgeries where surgeons can only rely on cameras. Computer vision is also beneficial in automatic depiction of tumors faster, thus helping doctors to start the treatment sooner [21-23].

B. Sports

The world of sports intrinsically involves fast and accurate motion which is challenging for competitors and master, as well as difficult for coaches and trainers to analyse, and for audiences to follow. The nature of most sports means that monitoring by the use of sensors or other devices fixed to players or equipment is generally not possible. Analysing or visualizing the motion of players and object detection at key moments in sport though computer vision can give useful insights to both coaches and broadcasters [24-28]. Training at elite level may involve the use of dedicated motion capture systems with multiple calibrated cameras and markers placed on the players. In elite games, computer vision is used to analyse individual player motion and can detect when players have become fatigued and require substitution. Apps are available for coaches of amateur athletes that can also improve performance and reduce the risk of injury (e.g., Coach's Eye). The sports events which are not televised, it becomes difficult for the performance analyst to gather the entire video from different angles. As a result, athletes must spend considerable amount of time in manually jotting down the events, hence computer vision has the potential to cover the existing the gap by helping in tracking specific player though object detection and image processing algorithms in a video [29]. Furthermore, the camera orientation, angles, position vary in different sport activities, however, computer vision algorithms deploying object detection and image processing helps in distinguishing between foreground and background objects. For instance, segmentation algorithms based on colour are presently used in sports to identify the grass as background in an image or video by its green colour, whereas players and objects are considered as foreground [30-32]. Computer Vision systems have been used in sport for refereeing support including cricket, tennis, and soccer (e.g., Hawkeye).

C. Education

The utilization of computer vision algorithms in education are beneficial to students as they help in maximizing students' academic output by offering a customized learning experience based on their strengths and weaknesses [33,34]. The computer vision also helps in improving teacher student interaction by understanding the behaviour of students. Furthermore, the behaviour of students also aids in enhancing the comfort level amongst students by grouping them within their specific groups of interest

[33,34]. Furthermore, computer vision helps in in understanding behaviour of students and interaction during diverse group tasks on how they teach others, and how comfortable they are with fellow students [35,36].

D. Agriculture

In agriculture deployment of computer vision made it possible for farmers to make yield estimation using their smartphones or tablets. One Soil Platform streamlines farming [37,38]. They develop solutions that help collect field data and monitor plants. More importantly, the technology can help perform routine and time-consuming tasks like planting, harvesting, and evaluating plant health and development [39,40]. All rolled into one, it does help farmers streamline their work. Furthermore, through camera in drones, computer vision helps farmers to analyse their crops easily, and also selectively sort crop by customer quality indicators (eg: shape/colour). Additionally, computer vision is also used by farmers to track, count and weight of the animals, thus saving a considerable amount of manual work. Computer Vision along with machine learning also beneficial and extensively used in agriculture due to the detection of several different crop diseases, enabling farmers to act on time and save a lot of work and money, and ensuring health and safety standard measures [41,42].

E. Robotics

In traditional robotics applications, the work environment must be fixed, and the positions are fixed, which means adapting the work environment to the robot through positioners etc. However, when the robots are combined with computer vision, they are flexible, less error prone, able to optimise their movements, improve the quality of the process and tasks. Robots are used in pick and place different parts and components using computer vison algorithms. In industrial robot applications computer vision extracts relevant information from images and analyses images to recognize the objects, including their size and colour [43-45]. After object detection and image processing, the information can be passed on to other systems to act. Some robots even have cameras embedded in their arms or heads to feed the information collected directly back to the robot [46]. Furthermore, with object detection and computer vision, robots can detect which parts should go in a particular piece and which parts to be excluded, thus determining the quality control of the product. With the extensive use of computer vision in robots, robots have found its existence in surgery too where they assist the doctors in invasive surgery through images and videos [47]. It is envisioned that in future with the help of computer algorithms robots will be able to perform of the tasks of humans in short span of time.

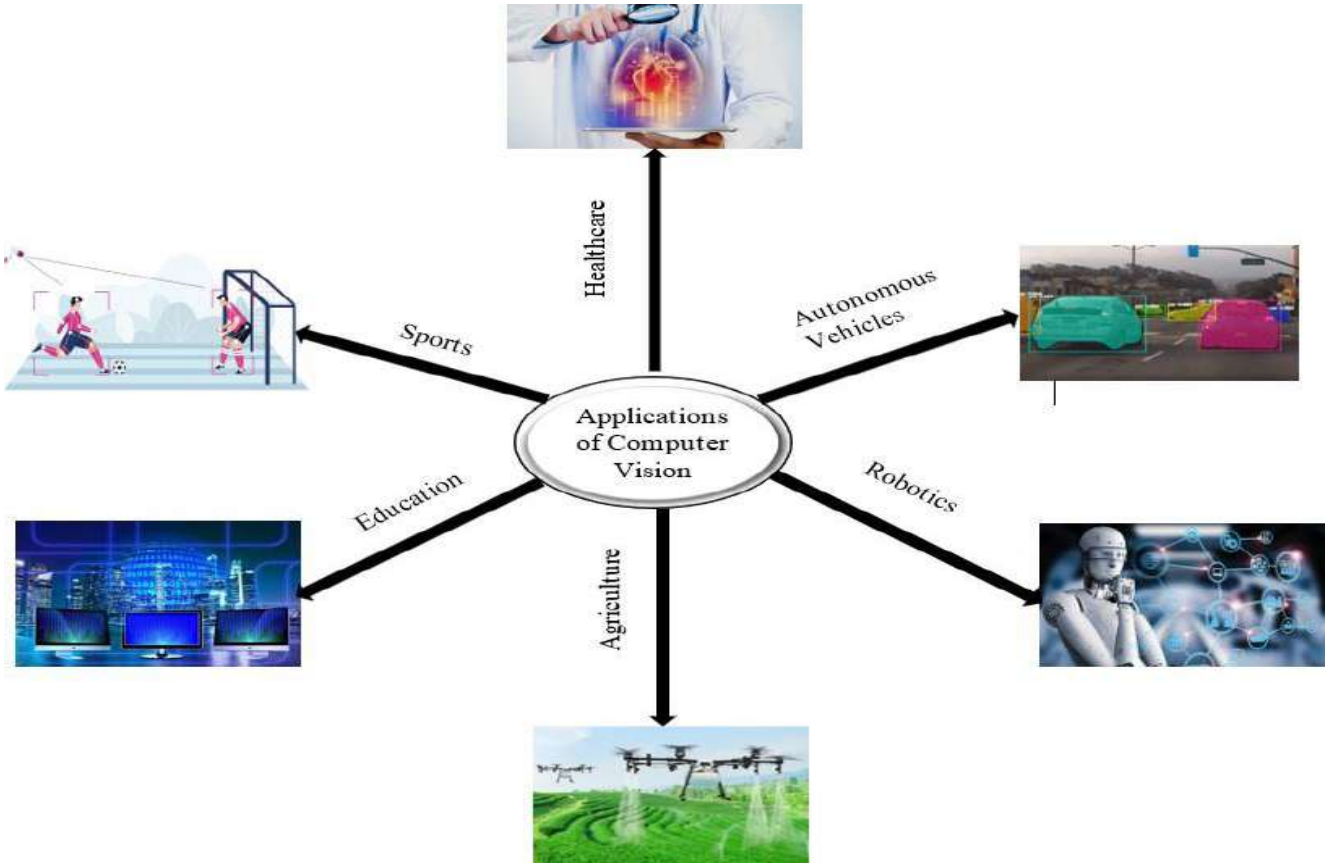


Fig.5 Applications of Computer Vision.

F. *Automated Vehicles*

In autonomous the computer vision plays a crucial role in detecting surrounding objects through object detection and image processing. The captured image and videos are converted and processed using algorithms that are easily read and understood by autonomous vehicles [46,47]. For example, detection of pedestrians, lanes, traffic signal, environment, is essential for automated vehicles to avoid collision and obstacles and all this is made possible through computer vision [48,49]. The camera deployed in autonomous vehicles capture the images and videos which are processed through computer vision system and the information is sent back to autonomous vehicle about the surroundings thus resulting in a successful working of autonomous vehicles [50,51]. The primary task of driving requires lateral and longitudinal control through a real-time understanding of the current situation, and therefore automated vehicles have a reliance on Computer Vision technologies.

IV CONCLUSIONS

Computer Vision (CV) is an interdisciplinary field that builds systems and analyses data from the images and videos captured by a camera through smart phone, laptop etc. This paper reviewed the several aspects and areas of computer vision. The basis of image processing, pixel, and images in context to computer were also explained in this paper. With the advancement of technology in recent years, computer vision has found its existence in numerous areas and have several applications, some of which were also discussed and mentioned in this paper. However, some of the challenges associated with computer vision technology and the future research directions are elaborated in next section of this paper.

V FUTURE RESEARCH OPPORTUNITIES

Computer Vision is an emerging technology with a vision that the computers in future would reduce the tedious manual work of several hours to minutes. Computer vision technology involves the use of image processing (IP), objection detection (OD) along with algorithms of artificial intelligence (AI) and machine learning (ML). Therefore, to train a system or model, there is a requirement of professional with complete understanding of computer vision. Furthermore, in certain applications of computer vision such as security, continuous monitoring is mandatory. Hence, in instances of technical issues, companies, organisations, or institutions working on computer vision are obliged to have a committed team of upskilled employees or team members. Additionally, challenge of reading pixel of images or videos at different orientations, angle, sizes in real-time situations in computer vision continue to exist. Despite these challenges, computer vision systems have enormous opportunities and capabilities for companies, organizations and institutions in future. Some of the future research directions of computer vision are discussed below:

A. *Training models based on real world situations*

In computer vision the complexity is to train the model that generalises the real-world situations because there is a possibility that the images and videos captured could be at different viewing angles, sizes, orientations, and shapes. Therefore, the model training becomes difficult especially in autonomous vehicles, robotics, and navigation where a small change in image or video parameters could result in an adverse event [45,46]. Hence, due to the vulnerability of the existing models, the requirement to develop models relating to real-world scenarios is an open area of research and future research direction.

B. *Online Education Algorithms*

In the situations such as the 2020-21 pandemic, computer learning or meta learning has gained significant attention. However, there is concern over online courses in terms of commercial protection, privacy, and opportunity for cheating. The ethics of placing students under continuous monitoring through camera or computer vision learning process is still an ongoing issue. Hence, stringent security measures need to be adopted by institutions implementing computer vision technology to protect user data from being misused [36-38]. The development of effective computer vision learning algorithms for student learning is an open area of research and future work.

C. *Security*

With the use of computer vision technology, it is possible to monitor and engage employees in online conferences or meetings through video conferencing, thus resulting in identifying the mood of a person. The deployment of computer vision for face detection, facial recognition, object detection in understanding the behaviour of a person is quite susceptible to errors such as both technical and non-technical. Therefore, decision making based on dependency of facial detection and recognition algorithms is a matter of concern for issues of privacy and security [22,23]. Hence, to provide the reliable and secure system, the security algorithms in computer vision free of any errors is a future work.

D. *Human Comfort*

The aviation industry is stepping towards innovative technologies to improve the human comfort, future aircraft will be designed differently to make them more sustainable. Human comfort perception is a key factor for aircraft manufacturers because willingness to use similar aircraft again for travelling is influenced by the human comfort. Real-time and effective human

discomfort detection plays a critical role in achieving energy efficient control vehicle environments. In long distance travel, the seat comfort is essential. Therefore, when designing seats, an indication of the perception of comfort/discomfort can be useful either for research and development purposes or potentially for automated systems to take actions that might mitigate discomfort amongst the passengers [52-54]. In [55], authors conducted an experimental research study of 60 minutes to analyse seat fidget movements (SFM) and demonstrated a computer vision system to detect the movements of the passenger associated with discomfort. The SFM model layout is shown in Fig 6, it is envisioned that the deployment of computer vision algorithms and analysing the captured video or images in future to detect the discomfort levels amongst passengers will significantly assist in enhance the comfort perception of passengers.

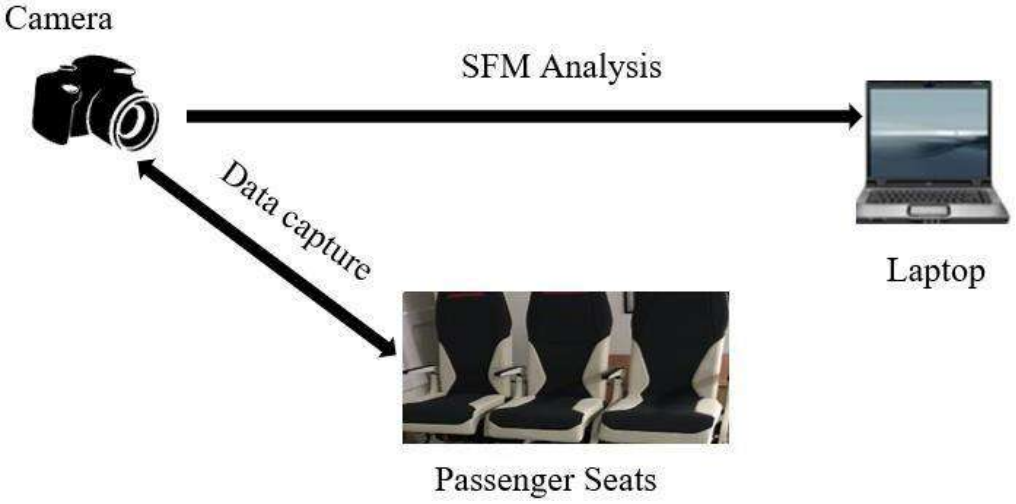


Fig.6 SFM model layout deploying computer vision using camera to capture data of passengers though video and anlaysis using software tools.

The field of computer vison has evolved in recent years with numerous applications in different areas deploying image processing and object detection algorithms. These algorithms will not only help in analysing the data but will also be beneficial to different industries such as aviation industry in optimising the interior of aircraft based on comfort perception of passengers.

ACKNOWLEDGEMENT

Partly supported by EU CleanSky ComfDemo — H2020- CS2-CFP08-2018-01.

REFERENCES

[1] Andrew Mort, AI in Customer Service, Posted on (Mar 11, 2020). How Computer Vision Applications are changing the World. https://techsee.me/blog/computer-vision

[2] Dana H. Ballard; Christopher M. Brown (1982). Computer Vision. Prentice Hall. ISBN 978-0-13-165316-0.

[3] Huang, T. (1996). Vandoni, Carlo, E (ed.). Computer Vision : Evolution And Promise (PDF). 19th CERN School of Computing. Geneva: CERN. pp. 21–25. doi:10.5170/CERN-1996-008.21. ISBN 978-9290830955.

[4] Milan Sonka; Vaclav Hlavac; Roger Boyle (2008). Image Processing, Analysis, and Machine Vision. Thomson. ISBN 978-0-495-08252-1.

[5] Reinhard Klette (2014). Concise Computer Vision. Springer. ISBN 978- 1-4471-6320-6.

[6] Linda G. Shapiro; George C. Stockman (2001). Computer Vision. Prentice Hall. ISBN 978-0-13-030796-5.

[7] Tim Morris (2004). Computer Vision and Image Processing. Palgrave Macmillan. ISBN 978-0-333-99451-1.

[8] Bernd Jähne; Horst Haußecker (2000). Computer Vision and Applications, A Guide for Students and Practitioners. Academic Press. ISBN 978-0-13-085198-7.

[9] David A. Forsyth; Jean Ponce (2003). Computer Vision, A Modern Approach. Prentice Hall. ISBN 978-0-13-085198-7.

[10] Kagami, Shingo (2010). 'High-speed vision systems and projectors for real-time perception of the world'. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops.IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops. 2010. pp. 100–107.

[11] Aggarwal, Geetika (2020) Visible Light Optical Camera Communication for Electroencephalography Applications. Doctoral thesis, Northumbria University.

[12] Richard Szeliski (30 September 2010). Computer Vision: Algorithms and Applications. Springer Science & Business Media. pp. 10–16. ISBN 978-1-84882-935-0.

[13] Papert, Seymour (1966-07-01). 'The Summer Vision Project'. MIT AI Memos (1959 - 2004). hdl:1721.1/6125.

[14] Margaret Ann Boden (2006). Mind as Machine: A History of Cognitive Science. Clarendon Press. p. 781. ISBN 978-0-19-954316-8.

[15] Takeo Kanade (6 December 2012). Three-Dimensional Machine Vision. Springer Science & Business Media. ISBN 978-1-4613-1981-8.

[16] Nicu Sebe; Ira Cohen; Ashutosh Garg; Thomas S. Huang (3 June 2005). Machine Learning in Computer Vision. Springer Science & Business Media. ISBN 978-1-4020- 3274-5.

[17] William Freeman; Pietro Perona; Bernhard Scholkopf (2008). 'Guest Editorial: Machine Learning for Computer Vision'. International Journal of Computer Vision. 77 (1): 1. doi:10.1007/s11263-008-0127-7. ISSN 1573-1405.

[18] E. Roy Davies (2005). Machine Vision: Theory, Algorithms, Practicalities. Morgan Kaufmann. ISBN 978-0-12-206093-9.

[19] Willie Brink. COMPUTER VISION, Applied Mathematics, Stellenbosch University, July 2013

[20] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins. Digital Image Processing Using MATLAB, Gatesmark Publishing,2nd, 2009

[21] Ferreira, P. M., Mendonça, T., Rozeira, J., & Rocha, P. (2012). An annotation tool for dermoscopic image segmentation. Proceedings of the 1st international work shop on visual interfaces for ground truth collection in computer vision applications - VIGTA '12, 1–6. Capri, Italy. https://doi.org/10.1145/2304496.2304501

[22] Balducci, F., & Borghi, G. (2017). An Annotation Tool for a Digital Library System of Epidermal Data. In C. Grana & L. Baraldi (Eds.), Digital Libraries and Archives (Vol. 733, pp. 173–186). Springer International Publishing. https://doi.org/10.1007/978-3-319-68130-6_14

[23] Mata, C., Lalande, A., Walker, P., Oliver, A., & Martí, J. (2017). Semi- automated labelling of medical images: ben efits of a collaborative work in the evaluation of prostate cancer in MRI. ArXiv:1708.08698 [Physics]. http://arxiv. org/abs/1708.08698

[24] S. Anderson. Forsgren helps revolutionize golf on TV with Protracer. Golf WRX, 11 July 2013.

[25] I. Atmosukarto, B. Ghanem, M. Saadalla, and N. Ahuja. Recognizing team forma tion in american football. In T.B. Moeslund, G. Thomas, and A. Hilton, editors, Computer Vision in Sports, chapter 13. Springer, 2014.

[26] S. K. Behendi, S. Morgan, and C. B. Fookes. Non-invasive performance measurement in combat sports. In P. Chung, A. Soltoggio, W. C. Dawson, Q. Meng, and M. Pain, ed itors, Proceedings of the 10th International Symposium on Computer Science in Sports (ISCSS), pages 3–10. Springer International Publishing, 2016.

[27] C. Bialik. The people tracking every touch, pass and tackle in the world cup. Five-Thirty Eight, 10 June 2014.

[28] A. Bialkowski, P. Lucey, P. Carr, S. Sridha ran, and I. Matthews. Representing team be haviours from noisy data using player role. In T.B. Moeslund, G. Thomas, and A. Hilton, editors, Computer Vision in Sports, chap ter 12. Springer, 2014.

[29] 46 S.Indu, Manjari Gupta and Prof. Asok Bhattacharyya.," Vehicle Tracking and Speed Estimation using Optical Flow Method". , International Journal of Engineering Science and Technology (IJEST)., Vol. 3 No. 1 Jan 2011

[30] J. Li, T. Wang, W. Hu, Y. Zhang, "Soccer Highlight Detection using Two Dependent Bayesian Network," In Proc. of IEEE International Conf. on Multimedia & Expo (ICME), 2006. [13] P. Viola, M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," In Proc. IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR), 2001.

[31] E. Wengrowski, W. Yuan, K. J. Dana, A. Ashok, M. Gruteser, N. Mandayam, Optimal radiometric calibration for camera-display communication, in: IEEE Winter Conf. on App. of Computer Vision, 2016, pp. 1–10. [165] M. Sonka, V. Hlavac, and R. Boyle, Image Processing, Analysis, and Machine Vision. Thomson-Engineering,2007.

[32] M. Kristan, J. Pers, M. Perse, and S. Ko vacic. Closed-world tracking of multiple inter acting targets for indoor-sports applications. Computer Vision and Image Understanding, 113(5):598–611, 2009.

[33] Wang A, Zhang W, Wei X. A review on weed detection using ground- based machine vision and image processing techniques. Comput Electron Agric 2019;158:226–40.

[34] Patrı´cio DI, Rieder R. Computer vision and artificial intelligence in precision agriculture for grain crops: a systematic review. Comput Electron Agric 2018;153:69–81.

[35] Ray PP. Internet of things for smart agriculture: technologies, practices and future direction. AIS 2017;9 (4):395–420. https://doi.org/10.3233/AIS-170440.

[36] Liakos KG, Busato P, Moshou D, et al. Machine learning in agriculture: a review. Sensors 2018;18:1–29.

[37] Rico-Ferna´ndez MP, Rios-Cabrera R, Castela´n M, et al. A contextualized approach for segmentation of foliage in different crop species. Comput Electron Agric 2019;156:378–86.

[38] Gomes JFS, Leta FR. Applications of computer vision techniques in the agriculture and food industry: a review. Eur Food Res Technol 2012;235:989–1000.

[39] Foglia MM, Reina G. Agricultural robot for radicchio harvesting. J Field Rob 2006;23:363–77.

[40] Mochida K, Koda S, Inoue K, et al. Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective. GigaScience 2019;8:1–53.

[41] Vazquez-Arellano M, Griepentrog HW, Reiser D, et al. 3-D imaging systems for agricultural applications – a review. Sensors 2016;16:1–24.

[42] J. Y. Lee and W. Hoff. Activity identification utilizing data mining techniques. In IEEE Workshop on Motion and Video Computing (WMVC), 2007.

[43] N.Mavridis, "A review of verbal and non-verbal human–robot interactive communication," Robotics and Autonomous Systems, vol. 63, pp. 22-35, 2015.

[44] G.-J. M. Kruijff, H. Zender, P. Jensfelt, and H. I. Christensen, "Clarification dialogues in human-augmented mapping," in Proc. of the 1st Annual Conference on Human Robot Interaction, HRI'06, (Salt Lake City, UT), Mar. 20

[45] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Recognition with cortex-like mechanisms," IEEE Transactions on Pattern Analysis and Machine Intelli gence, vol. 29, no. 3, pp. 411– 426, 2007.

[46] G. Hager and K. Toyama, "The XVision system: A general–purpose substrate for portable real–time vision applications," Computer Vision and Image Understanding, vol. 69, no. 1, pp. 23–37, 1996.

[47] R. T. amd D. Stoianovici, "Medical robotics in computer-integrated surgery," IEEE Transactions on Robotics and Automation, vol. 19(5), pp. 765– 781, 2003.

[48] Grigorescu, S., Trasnea, B., Cocias, T., & Macesanu, G. (2020). A survey of deep learning techniques for autono mous driving. Journal of Field Robotics, 37(3), 362–386. doi: 10.1002/rob.21918.

[49] Lladó, X., Mahmood, M. H., & Salvi, J. (2016). Semi automatic tool for motion annotation on complex video sequences. Electronics Letters, 52(8),

602–604. doi: 10.1049/el.2015.4163.

[50] Schöning, J., Faion, P., & Heidemann, G. (2015). Semi automatic ground truth annotation in videos: an inter active tool for polygon- based object annotation and segmentation. Proceedings of the knowledge capture con ference on ZZZ - K-CAP 2015, 1–4. Palisades, NY, USA. https://doi.org/10.1145/2815833.2816947

[51] Wang, B.-L., King, C.-T., & Chu, H.-K. (2018). A semi-automatic video labeling tool for autonomous driv ing based on multi-object detector and tracker. 2018 sixth international symposium on computing and networking (CANDAR), 201–206. Takayama, Japan. https://doi.org/ 10.1109/CANDAR.2018.00035.

[52] Varela, M., Gyi, D., Mansfield, N., Picton, R., & Hirao, A. (2017). Designing movement into automotive seating - does it improve comfort?

[53] Neil Mansfield, Alessandro Naddeo, Susanne Frohriep, Peter Vink, Integrating and applying models of comfort,Applied Ergonomic s,Volume 82,2020,102917,ISSN 0003-6870, https://doi.org/10.1016/j.apergo.2019.102917

[54] Mansfield, N., Sammonds, G.M., Darwazeh, N., Massoud, S., Mocio, A., Patel, T., & Sehdev, A. (2017). Movement analysis to ind icate discomfort in vehicle seats.

[55] Sammonds, George; Fray, Michael; Mansfield, Neil J. (2017): Effect of long term driving on driver discomfort and its relationship with seat fidgets and movements (SFMs). Loughborough University. https://hdl.handle.net/2134/22441

Walkthrough to Convolution Neural Network

Akshita Sharma[1], Sumant Bansal[2], Rohit Raj[3], Shimpy Goel[4]

**Student, Dept. Of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology and Management New Delhi, India*

[1]akshita.2507.sharma@gmail.com

[2]bansalsumant7@gmail.com

[3]rajr26010@gmail.com

[4]Shimpy Goyal

Assistant Profesor, Dept. Of Information Technology, Dr Akhilesh Das Gupta Institute of Technology and Management New Delhi, India

[4]shimpygoyaluiet@gmail.com

***Abstract*: This paper is based on the information related to Convolution Neural Network which is the backbone of modern deep learning techniques. Here we discuss this network and also highlight the visual problems. CNN is built on prior knowledge of human vision systems. The primary focus of this paper is to learn different models in Convolution Neural Network and challenges to solve different vision problems. Further, we will focus on various milestones in the ImageNet challenge. Many of these basic ideas and concepts are still studied by many researchers as any new concept always originates from these ideas. Later we will highlight major models of CNN like AlexNet, VGG, Residual Net, GoogLeNet. Most importantly, this paper provides a brief knowledge of Convolution Neural Network with its resemblance to the Human Vision System. And in the end, we will conclude the learning and usage of this network in modern deep learning techniques.**

***Keywords*- Convolutional Neural Network, Vision Problem, Visual Cortex, CNN, Neocognitron**

I. INTRODUCTION

In this paper, we will discuss a bit different family of models i.e, about the Convolutional Neural Network (CNN) family. This CNN family mainly evolves from the knowledge of the human visual cortex. The most important reason that is responsible for the success of convolutional neural networks in vision problems is because of its bionic design to replicate or simulate human vision systems.

A. Visual Cortex

Convolutional Neural Network is widely known as being inspired by the visual cortex, however, except that some publications discuss this inspiration briefly (Poggio and Serre, 2013[2]; Cox and Dean, 2014)[3], few resources present this inspiration thoroughly.

The visual cortex of the brain which is located at the back of the skull is a major part of the cerebral cortex that plays an important role in processing visual information.

Visual information coming from the eye passes through a series of brain neurons and reaches the visual cortex. The parts of the visual cortex that receive the sensory inputs are known as the primary visual cortex, which is also known as area V1. Visual information is further managed by extrastriate areas of the brain that includes visual areas two (V2) and four (V4). There are also three other visual areas (V3, V5, and V6), but here, we primarily focus on the visual areas that are related to object recognition, which is also called ventral stream. It also consists of Inferior temporal gyrus which is one of the higher levels of the ventral stream of visual processing, associated with the representation of complex object features, such as global shape, like face perception.

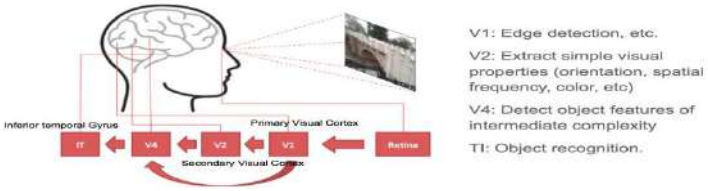


Fig. 1 Illustration of Ventral Stream of Visual Cortex

B. Neocognitron and Visual Cortex

It is a neural network that consists of two different kinds of layers, S-layer as a feature extractor and C-layer as structured

connections to organize the extracted features.
S-layer consists of several S-cells that are inspired by the cell in the primary visual cortex. It serves as a feature extractor. Each S-cell can be ideally trained to be responsive to a particular feature presented in its respective field. Usually, local features such as edges in particular orientations are extracted in lower layers while global features are extracted in higher layers. This structure highly resembles how humans conceive objects. The C-layer resembles complex cells in the higher pathway of the visual cortex. It is mainly introduced for the shift-invariant property of features extracted by S-layer.

a. *Parameter Learning*

During the parameter learning process, only the parameters of S-layer are updated. Neocogitron can also be trained but unsupervised, for a good feature extractor out of S-layers. The training process for S-layer is very similar to the Hebbian Learning rule, which strengthens the connections between S-layer and C-layer. This training mechanism also introduces the problem that the Hebbian Learning rule introduces, which is the strength of connections will saturate, as it keeps increasing.

C. *CNN and Visual Cortex*

Now we have proceeded from Neocogitron to the Convolutional Neural Network. In this section, we will first introduce the building components: convolutional layer and subsampling layer. Then we group these components to represent Convolutional Neural Network by using LeNet as an example.

a. *Convolution Operations*

The convolution operation is simply just a mathematical operation, which should be treated equally with other operations like addition or multiplication and should not be discussed particularly in machine learning literature. However, we still discuss it here for the continuation of the topic.
Let convolution be a mathematical operation on two functions f and g and produce a third function h, which is integral that expresses the amount of overlap of one function (f) as it is shifted over the other function (g). It is denoted by h = f * g and is described as:

$$h(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$$

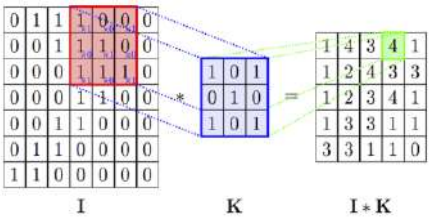


Fig. 2 A simple example of 2D convolution

One should know that convolution is locally shifting invariant, which means that for many different combinations of how the nine numbers in the upper 3 × 3 matrices are placed, the convoluted result will be the same. This fixed and unchanging property plays a major role in vision problems because in an ideal case, the recognition result should not be changed due to any shift or rotation of features.

b. *Connections between CNN & Visual Cortex*

With the ideas about two-dimensional convolution, we further discuss how convolution is a useful operation that can replicate the tasks performed by visual cortex.
The convolution operation is also known as kernels. By different choices of kernels, different operations of the images could be achieved. These operations typically include identity, edge detection, blur, sharpening etc. By introducing random matrices as convolution operators, some interesting properties might be discovered.

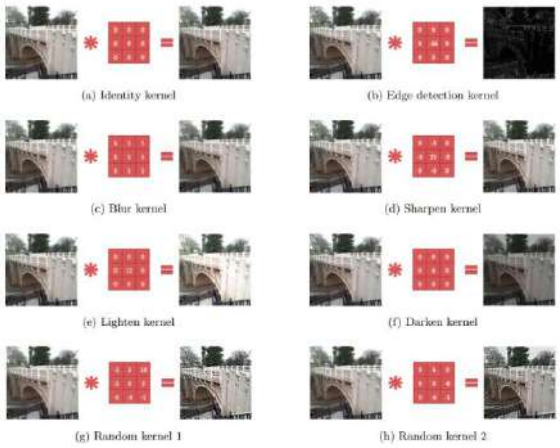


Fig. 3 Examples of Kernels

The above figure is an illustration of some example kernels that are applied to the same figure. One can see that different kernels can be applied to fulfil different tasks. Random kernels can also be applied to transform the image into some interesting results.

D. LeNet: The Pioneer of CNN

LeNet is devoted as a model that is widely recognized as the first convolutional neural network.

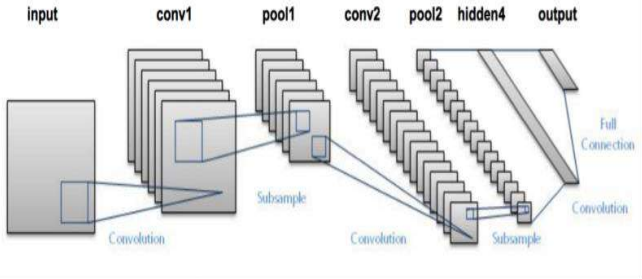


Fig. 4 Illustration of LeNet

The above figure shows an illustration of the architecture of LeNet. It consists of two pairs of Convolutional Layer and Subsampling Layer (Pooling Layer) and is further connected with a fully connected layer and an RBF layer for classification.

a. Convolutional Layer

A convolutional layer, as the name suggests, is a layer that performs convolution operation. As we have discussed in the previous section, a clever selection of convolution kernels can effectively replicate the task of the visual cortex. Convolutional layer introduces another operation, the non-linearity transforms after convolution to assist the simulation to be more successful.
Non-linearity transform helps to remove the negative part of the input, resulting in a clearer contrast of meaningful features and can be described as:
$f(x) = \max(0, x)$

b. Sub-Sampling Layer

Subsampling Layer samples out input from every region it looks into. Some different strategies of sampling can be considered, like max-pooling, average-pooling or even probabilistic pooling i.e, taking a random one.
Sampling helps to turn the input representations into smaller and more manageable embeddings. It also makes the network differ from small transformations, distortions, and translations in the input image. A small distortion/disturbance in the input will not change the outcome of pooling since we take the maximum or average value in a local neighbourhood.

E. Milestones in ImageNet Challenge

With the success of LeNet, Convolutional Neural Network has shown great potential in solving vision tasks. Along with this, several great milestones have attracted huge attention and have been applied to other fields with good performance. In this section,

we will briefly discuss some of these models.

a. *AlexNet*

While LeNet is the major contribution to the era of convolutional neural networks, AlexNet which was invented by Krizhevsky in 2012[8] becomes one that starts of the era of CNN used for ImageNet classification. AlexNet is the first evidence that CNN can perform well on this ImageNet dataset and it performs well so that it leads the society into a competition of developing CNNs. The success of AlexNet is not only due to its unique design of architecture but also due to the clever mechanism of training. To avoid the expensive training process of this clever mechanism, AlexNet has been split into two streams and is trained on two GPUs. It also used data augmentation techniques that consist of image translations, horizontal reflections, and patch extractions, etc.

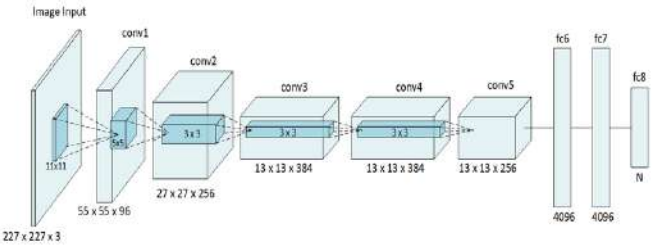


Fig. 5 Illustration of AlexNet

Illustration of the architecture of AlexNet is shown in the above figure. However, there is not much to learn from this architecture, although of its remarkable performance. This architecture doesn't even have a well-supported theory to explain which causes many researchers blindly burn computer memory to form new proofed architecture.

b. *VGG*

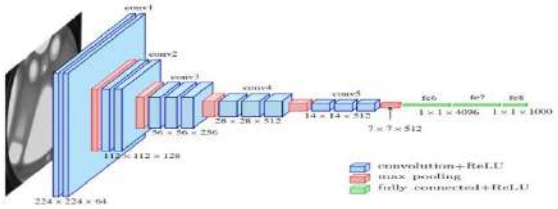


Fig. 6 Illustration Of VGG

VGG is a deeper architecture with 19 layers than other models around that time, the architecture is extremely simplified. Here, all the layers are 3 × 3 convolutional layers with a 2 × 2 pooling layer. This simple usage of convolutional layer simulates a larger filter while keeping the benefits of smaller filter sizes because the combination of two 3×3 convolutional layers has an effective responsive field of a 5 × 5 convolutional layer, but with fewer parameters.

The occupying space of the input volumes at each layer will decrease as a result of the convolutional and pooling layers, but the depth of the volumes increases because of the increased number of filters. This behaviour forces the idea of VGG to shrink spatial dimensions, but grow depth subsequently.

c. *Residual Net*

Residual Net is also known as ResNet and is a 152 layer network, which was almost ten times deeper than previous popular architecture, VGG. ResNet explores deeper structure with simple layers. However, by increasing the number of layers will result in worse results, for both training cases and test cases.

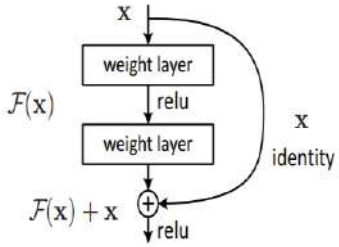


Fig. 7 Illustration of Residual Net

ResNet is extended to a 1000-layer version with success on CIFAR data set.
The property of its shallow network allows ResNet to perform as a collection of independent networks; each network is significantly shallower than the integrated ResNet itself. This property also explains why gradients can be passed through the ultra-deep architecture without being diminished.

d. *GoogLeNet Model*

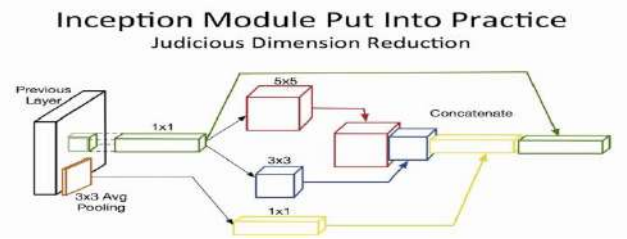


Fig. 8 Illustration of GoogLeNet

GoogleNet is also known as Inception V1 and was proposed by Google itself with the help of contributions from various universities in 2014[14]. This architecture was the winner at the ILSVRC 2014[15] image classification challenge. It has provided a significant decrease in error rate as compared to previous winners AlexNet and VGG.

i. Features of GoogLeNet:

The GoogLeNet architecture is very different from previous state-of-the-art architectures such as AlexNet and VGG. It uses many different kinds of methods such as *1×1* convolution and global average pooling that enables it to create deeper architecture and thus produce better results. Some features are described below:

- **1×1 convolution:** This architecture uses 1×1 convolution in its architecture. These convolutions used to decrease the number of parameters.
- **Global Average Pooling:** This method is used at the end of the network. This layer takes a feature map of *7×7* and averages it to *1×1* which makes it to decrease the number of trainable parameters to 0 and also improves the accuracy by 0.6%
- **Auxiliary Classifier for Training**: This architecture uses some intermediate classifier branches in the middle of the architecture, which is used during training only. Thus, these layers help in combating gradient vanishing problems and also provide regularization.
- **Inception Module:**In this architecture, the convolution size for each layer is fixed. In the Inception module convolution and max-pooling are performed in a parallel way and at the input and the output of these are stacked together to generate the final output.

F. *Challenges and Chances for Fundamental Vision Problems*

ResNet is not the end, new models and techniques appear every day to push the limit of CNNs further. For example, Zhang took a step further and put Residual Block inside Residual Block.

a. *Network Property and Vision Blindness Spot*

Convolutional Neural Networks have reached an unmatched accuracy in object detection. However, it may still be far from industry reliable application due to some fascinating properties found by Szegedy in 2013[12].

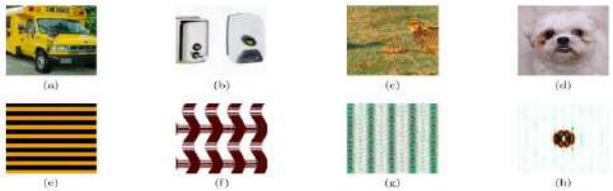


Fig. 9 Examples from Work done by Szegedy

The above figure shows some examples that are based on work done by Szegedy. Here images, (a) - (d) are distinguishable but the difference between them and original cannot be observed by the human eye, while images (e) - (h) are not distinguishable but the neural network predicts these images correctly.
Even without these examples, one may also observe that the reliable prediction of neural networks could be an issue due to the fundamental property of a matrix that there is the existence of null space. Nullspace acts like a blind spot to a matrix and changes within null space are never sensible to the corresponding matrix.

b. *Human Labeling Preference*

Fig. 10 Examples of Human Label Preference

The actual error rate of ResNet is 3.6% which might be called false after this result. For example, fig (a), (b), (h), are considered as small parts of the image, while there are many contrasting things expressed by image. While fig (d), (e) are considered as background images.

II. Literature Survey

As CNN resembles the human vision system, we can easily introduce the successful models that make themselves famous through the ImageNet Challenge done by Jia Deng in 2009[1].
Hubel and Wiesel, 1959[4], lays the ground for further study in Convolutional Neural Networks.
Neocogitron was first introduced by Fukushima in 1980[5], which are generally seen as the model that inspires Convolutional Neural Networks on the computation basis.
The solution to problem that, the strength of connections will saturate, as it keeps increasing introduced by Fukushima in 1980[5], which were introduced with the name “inhibitory cell”. It performed the function as normalization to avoid this problem.
LeNet was introduced by LeCun in 1990[6] and is inspired by the Neocogitron.
According to Russakovsky, 2015[7], the potential for solving vision tasks has attracted a large number of researchers regarding object recognition in CIFAR classification and ImageNetclassification.
In the competition of exploring different and complex architectures, Simonyan and Zisserman in 2014[9] showed that simplicity will be a favourable direction with a model named VGG.
Kaiming He in 2016[10], verified that residual blocks are essential for travelling or propagating the information smoothly, therefore it simplifies the optimization.
Viet, et. All, 2016[16] showed that ResNet behaves like a group of shallow networks.
Zagoruyko and Komodakis in 2016[11] attempted to decrease the depth of the ResNet by increasing the width.
Szegedy in 2013[12]showed that they could force a deep learning model to misclassify an image simply by adding disturbance to that image. They have also shown that this property is more likely to be a modelling problem, in contrast to problems raised by insufficient training.

III. Future Scope

Although the recent advancements of deep learning have been impressive, there still exist challenges to its application to many fields.
ResNet is still not completely vacant from clever design architectures. The number of layers in this network and the number of layers that Residual Block allows identity to cross are still choices that require experimental validations. Nonetheless, to some level, ResNet has become a solution in the development of CNN, which is better than blind burning of computer resources. Also, the idea of Residual Block has been found in the actual visual cortex.
Wandell, 1995[13] shown that blind spots also exist in the human vision system which links this work with the flaws of the vision system to the defects in neural networks, which makes it a helping hand to overcome these defects in the future.
With the result of ResNet, new data sets are being introduced day by day like COCO in 2014, Flickr in 2015, and VisualGenome in 2016 may open a new era of vision problems with more competitive challenges. However, the basic or foundational problems and experiences introduced by ResNet should never be forgotten and must overcome in future.

IV. Conclusion

In this paper, we recall one of the primary concepts, Convolution Neural Network to practice modern deep learning with some challenges and changes in Vision problems. Further, we discussed some of the milestones in ImageNet Challenge like AlexNet, VGG, Residual Net, GoogLeNet etc. We also see the base model, LeNet which is the pioneer of CNN and has two layers.
This paper serves two goals:

1. It records the major developments in the Convolution Neural Network, with its resemblance with the Human Visual Cortex System. These similarities define a whole new branch of learning in the field of Deep Learning for researchers and students.
2. This paper should be able to suggest to the readers how these remarkable works are major contributors and how these are developed among thousands of other researches in this field.

We hope this helps the readers for a better understanding of CNN and its importance in current and future developments in Deep Learning techniques.

V. TABLE 1. COMPARISON TABLE

Discovered By:	**Model Name**	**Layers**	**Accuracy**
Kunihiko Fukushima	Neocognitron	s-layer & c-layer	52.50%
YannLeCun	LeNet	pairs of s & c-layers with connection Layer and RBF Layer	99.557%
Alex Krizhevsky	AlexNet	8 Layers	84.7%
Karen Simonyan& Andrew Zisserman	VGG	19 Layers	62.8%
Kaiming He	ResidualNet	152 Layers	69.7%
Google Inc	GoogLeNet	22 Layers	63%

REFERENCES

[1] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "Imagenet: A large scale hierarchical image database", 2009 IEEE Conference on Computer Vision and Pattern Recognition 2009, IEEE 2009, DOI: 10.1109/CVPR.2009.5206848.
[2]TomasoPoggio and Thomas Serre,"Models of the visual cortex", Scholarpedia, 2013, DOI: 10.4249/scholarpedia.3516
[3] David Daniel Cox and Thomas Dean,"Neural networks and neuroscience-inspired computer vision", Current Biology, 2014, DOI: 10.1016/j.cub.2014.08.026.
[4] David H Hubel and Torsten N Wiesel,"Receptive fields of single neurons in the cat's striate cortex",The Journal of Physiology, 1959, DOI: 10.1113/jphysiol.1959.sp006308.
[5] Kunihiko Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", Biological Cybernetics, 1980, DOI: 10.1007/BF00344251
[6] Le Cun, B Boser John S Denker, D Henderson, Richard E Howard, W Hubbard, and Lawrence D Jackel."Handwritten digit recognition with a back-propagation network", In Advances in neural information processing systems, Citeseer, 1990, DOI: 10.5555/2969830.2969879
[7] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, SanjeevSatheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, AdityaKhosla, Michael Bernstein, et al,"Imagenet large scale visual recognition challenge. International Journal of Computer Vision", 2015, DOI: 10.1007/s11263-015-0816-y.
[8] Alex Krizhevsky, IlyaSutskever, and Geoffrey E Hinton,"Imagenet classification with deep convolutional neural networks", In Advances in neural information processing systems, pages, 2012, DOI: 10.5555/2999134.2999257.
[9] Karen Simonyan and Andrew Zisserman,"Very deep convolutional networks for large-scale image recognition", 2014, arXiv: 1409.1556.
[10] Kaiming He, Xiangyu Zhang, ShaoqingRen, and Jian Sun,"Identity mappings in deep residual networks", 2016, arXiv: 1603.05027v3
[11] Sergey Zagoruyko and Nikos Komodakis,"Wide residual networks", 2016, arXiv: 1605.07146.
[12] Christian Szegedy, WojciechZaremba, IlyaSutskever, Joan Bruna, DumitruErhan, Ian Goodfellow, and Rob Fergus, "Intriguing properties of neural

networks", 2013, arXiv:1312.6199
[13] Brian A. Wandell,"Foundations of vision", Sinauer Associates, 1995.
[14] Google Inc, "Going Deeper with Convolutions", 2014, arXiv: 1409.4842v1.
[15] ILSVRC, "ImageNet Large Scale Visual Recognition Competition", 2014, ImageNet Large Scale Visual Recognition Competition (ILSVRC) (image-net.org)
[16] Andreas Veit, Michael J Wilber, and Serge Belongie, "Residual networks behave like ensembles of relatively shallow networks" In Advances in Neural Information Processing Systems, 2016, arXiv:1605.06431.

Synthetic Minority Over Sampling Technique for Class Imbalance Problem

Arushi Gupta
Assistant Professor
Department of lnformation Technology, ADGITM, Delhi, India.
arushi10818@gmail.com

***Abstract*- In Data-Level Approaches, the data-sets are balanced before applying TCT so that results may not get overwhelmed by the majority class. Data level techniques are further classified as undersampling, oversampling and hybrid approaches. The oversampling techniques increase the number of data points in the minority class either randomly through replication of the existing data or generating synthetically by using any technique to improve the imbalance ratio so that TCT can be used to classify the data. Under SMOTE (Synthetic Minority Over Sampling Technique), data points are synthetically generated in the minority class rather than replication. The paper includes techniques proposed by researchers to handle this problem using oversampling approach are also addressed. Also, find out the gaps in the existing techniques used for class imbalance learning techniques and the ways through which the research on class imbalance problem can be enhanced keeping in view the real-life problems.**

***Keywords*- Classification, Class imbalance problem, Oversampling, SMOTE, Hybrid sampling.**

I. INTRODUCTION

The techniques proposed by researchers to solve CIP are majorly classified into data-level approaches (Pre-processing techniques), algorithm level approaches and the hybrid level approaches or ensemble learning methods [1]. In data-level approaches, data-sets are balanced before applying TCT so that results may not get overwhelmed by the majority class. The algorithms that come under data level approach are based upon undersampling, oversampling and hybrid sampling [2]. In Algorithm Level Approaches, the internal algorithm structure is modified to improve the sensitivity of algorithm so that results of classification algorithms should not deviate towards the majority class. Third approach is the Hybrid Ensembles, which is the combination of data- level and algorithm level approaches[1,3].The algorithms that come under hybrid ensembles approach are based upon Rotation forest, bagging, boosting and also based on both bagging and boosting[4].

Data level techniques are those wherein data is pre-processed, before classifying the data, using traditional classifiers. The main aim of data level techniques is to rebalance the data before classification. The root cause of CIP is imbalanced data-sets that is one class has more number of instances than the other and the result of smaller class is influenced by the bigger class [1]. Moreover, the risk factor is directly proportional to the imbalance ratio that is as the datasets become more imbalance, the impact of majority class is more on the results of minority class. These techniques are also referred as external or pre-processing techniques. Data level techniques are further classified as undersampling, oversampling and hybrid approaches.

In the Undersampling approach, the working area in the data-set is majority class wherein data points are removed from the majority class either randomly or by using some technique to improve the imbalance ratio before classifying the data. Major advantage of undersampling approach over oversampling is that it is fast and simple [5]. But, in case of binary or multiclass detection, it may lead to the loss of informational data contained in the majority class. Oversampling techniques increase the number of data points in the minority class either randomly through replication of the existing data or generating synthetically by using any technique to improve the imbalance ratio so that TCT can be used to classify the data. Major advantage of oversampling approach is that no data is lost during the process as no data is removed, although new information is generated for rebalancing [6]. The limitation of oversampling is that it is computationally expensive as compared to undersampling approach because in this approach, data is generated. The overfitting problem may also occur as it will replicate the same data points to rebalance the data-set. Oversampling changed the minority class data by adding more instances, which in turn changed the actual data in the class that is not acceptable in some critical real time problems [1,7]. All those techniques that uses both oversampling as well as undersampling approach to improve the imbalance ratio before classification process comes under Hybrid Sampling techniques [7]

SMOTE[8] is an oversampling technique proposed by N. Chawla et al. in 2002. In SMOTE, data points are synthetically

generated in the minority class rather than replication. In this method, oversampling is done by considering all the minority class data and new data is generated by using interpolation method [8]. The oversampling rate required decides the number of neighbor to be chosen randomly from k- nearest neighbor. One of the main limitation is that SMOTE blindly generalizes data points in the minority class without checking the status of points as good points or noise.

The existing techniques of data level approaches described in the paper are SMOTE (Synthetic Minority Oversampling Technique (SMOTE), Borderline SMOTE (BorSMOTE), Safe-Level SMOTE (SL- SMOTE), Cluster SMOTE (C-SMOTE)), SMOTE+ Tomek link (SMOTE+TL), TL+C-SMOTE, SL-SMOTE+TL and BorSMOTE+TL. In this paper, Section II includes Oversampling Techniques, Section III describes Hybrid sampling techniques, Section IV acknowledges conclusion and future scope of the paper.

II. OVERSAMPLING TECHNIQUES

It increases the number of data points in the minority class either randomly through replication of the existing data or generating synthetically by using any technique to improve the imbalance ratio so that TCT can be used to classify the data [9].

A. SMOTE (Synthetic Minority Oversampling Technique)

SMOTE[8] is an oversampling technique proposed by N. Chawla et al in 2002. In SMOTE, data points are synthetically generated in the minority class rather than replication. In this method, oversampling is done by considering all the minority class data and new data is generated by using interpolation method. Thus, synthetic instances are generated in this approach. The need of synthetic generation is described by the figure 1. In the figure, Receiver Operating Characteristic (ROC) curve is shown that shows the performance of classifier by plotting between the true positive rate and false positive rate. The area under the curve is also used as a measure of identifying and choosing the optimal classifier [10]. The oversampling rate required decides the number of neighbors to be chosen randomly from k- nearest neighbor.

The k-nearest neighbor algorithm uses Euclidean distance between two points or tuples as a distance metric, say, $X_1 = (x_{11}, x_{12},..., x_{1n})$ and $X_2 = (x_{21}, x_{22},..., x_{2n})$, then Euclidean distance [1] is calculated as

$$dist(X_1, X_2) = \sqrt{\sum_{i=1}^{n} (x_{1i} - x_{2i})^2}$$

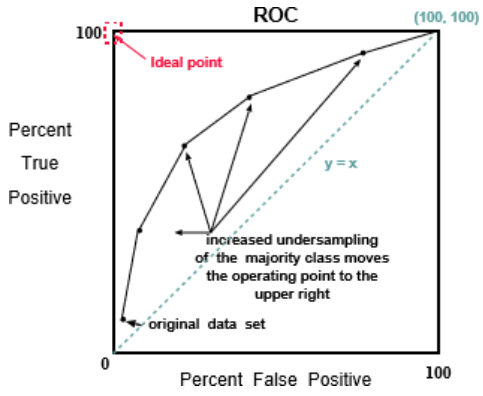


Fig.1 ROC curve sweeps out as performance moves from lower left point to upper right point through increase in undersampling [8]

To oversample the minority class instances, each instance/sample and synthetic instance is generated along the line segments that join any or all of k nearest neighbors minority class samples. Usually for the implementation, five nearest neighbors are used. One of the main limitation is that SMOTE blindly generalizes data points in the minority class without checking the status of points as good points or noise.

For generating synthetic instance, follow the steps:

1. Calculate difference between sample (feature vector) under consideration and its nearest neighbor.
2. Multiply that difference by random number that lies between 0 and 1, and then add it to sample/feature vector under consideration.

B. Cluster SMOTE

Cluster SMOTE [11] proposed by David A Cieslak et. al. in 2006, is a technique that uses K-means clustering algorithm [12] to find clusters of minority class examples and then applying SMOTE to oversample it. Such that the class imbalance ratio can be improved and both classes will have approximately the same number of instances for better prediction results by the classifier. Cluster-SMOTE applies unsupervised learning mechanism so that it will partition datasets into regions or clusters that will enable SMOTE to output better enhanced and accurate results. Thus, these results can be used as a proof for justifying the need for improvement over SMOTE.

The approach of this method for solving the class imbalance problem is that having a small group of minority examples makes difficult to establish proper class borders. Thus, the ability to correctly define the class regions and hence their borders would allow for trivial classification. As these regions are unknown and even in the best cases may be impossible to deduce from given data, only an approximation of these regions may be inferred. Even approximations may enhance classifier construction.

To develop these minority region approximations, apply simple k-means clustering to the set of minority examples in each dataset. After that, apply SMOTE to each cluster and then reform the dataset by reinserting the set of original minority examples and the synthetic examples as well. This process allows for focused improvements on a localization basis for the minority class and should improve SMOTE's performance on imbalanced datasets.

Clustering enable the detection of such distinct patterns and that generating synthetic examples focusing on localizations will enhance global classification. Cluster- SMOTE provides an improvement on SMOTE for the packets dataset, but cannot be used on destinations due to the limited feature space size of the minority class.

C. SAFE-LEVEL SMOTE

Safe-Level SMOTE [13] is an oversampling technique proposed by C. Bunkhumpornpat et al. in 2009. Safe-Level SMOTE (SL-SMOTE) is an extension of SMOTE, which assigns a safe level to every positive point before synthetically generating data points in the minority class. Every point, the algorithm synthetically generated lies near to the largest safe level as to assure that synthetically generated points lie in the safe regions only. Safe levels are defined by k-nearest neighbor. Any sample, whose safe level value is closer to zero, is considered as noise. Safe ratios are defined as safe level of a positive instance (p)/safe level of nearest neighbors (n). It helps in determining the safe positions for the generation of synthetic instances.

Five cases occur according to the value of safe level ratio:

First case happens as a result of safe level ratio value is ∞ and safe level of p is 0. Thus, it represents that both p and n are noises. Then, if it occurs, generation of synthetic instance will not occur as emphasization on noise regions is not important.

Second case happens when safe level ratio value is ∞ and safe level of p is not equal to 0. Thus, only n is noise. If it occurs, then generation of synthetic instance occurs from p through replication of p and noise instance n is avoided.

Third case happens when safe level ratio value is equal to 1. It means that safe level of p and n are same. If it occurs, then generation of synthetic instance is along line segment joining p and n because both p and n are safe.

Fourth case happens when safe level ratio value is greater than 1. It means that safe level of p is more than safe level of n. If it occurs, then generation of synthetic instance is closer to p in the range of [0, 1 / safe level ratio] as p is more safe than n.

Fifth case occurs when safe level ratio value is less than 1. It means that safe level of p is less than safe level of n. If it occurs, then generation of synthetic instance is closer to n in the range of [1 - safe level ratio, 1] as n is more safe than p in this case.

After each for loop iteration is completed, except first case, generation of synthetic instance is within range specified line segment between p and n. After that s is added to the set of all synthetic instances D'.

The technique produces |D| - t synthetic instances where number of positive instances in D is represented as |D|, and number of instances that fulfill condition for first case is represented by t.

D. BORDERLINE SMOTE

Borderline SMOTE [14] proposed by Han et. al in 2005, is based upon the concept that the data points which are close to or on

the borderline can be misclassified with greater than those which are far from the borderline thus are more important for classification. Thus, this method over-sampled only the borderline samples of minority class. It helps in achieving better prediction values by learning the borderline of each class in dataset as exactly as possible in training process. Borderline-SMOTE is different from other approaches as it does not oversample all minority examples or a random subset of these minority class data. Firstly, find the borderline minority examples; then, generation of synthetic minority examples from them, which are later, combined to the original training data.

Consider the training data is represented by T, the minority class data is represented by P and the majority class data is represented by N, and where, P = {p1, p2 ,..., ppnum}, N = {n1, n2,…., nnnum} ,also pnum and nnum represents the total number of minority and majority examples in the training dataset. The procedure for performing borderline-SMOTE is described below:

Step-1: Calculate m nearest neighbors from the provided dataset T, for each instance pi such that i=1,2,….,pnum of the minority class P. Then, among the m nearest neighbors obtained, determine the number of majority class examples m' ($0 \leq m' \leq m$).

Step-2: If all the m nearest neighbors are majority examples, i.e., m=m', then that instance pi of minority class is considered to be noise and thus discarded. Also, if $0 \leq m' < m/2$ is the case, then the instance pi is safe and thus also not considered for further steps.

Otherwise, if the case is $m/2 \leq m' < m$, then the instance pi is considered as DANGER as the number of majority examples in its nearest neighbor is less than majority examples which can be easily misclassified.

Step-3: The borderline data of minority class P are the examples considered as DANGER such that DANGER $\subset$P. Now, set DANGER = {p'1, p'2 ,..., p'dnum}, $0 \leq dnum \leq pnum$. Calculate k nearest neighbors from P, for every DANGER example.

Step-4: Now, synthetic positive examples that are dnum×s are generated from the data in DANGER, where s represents integer between 1 and k and for each p'i , s is number of nearest neighbors selected randomly from its k nearest neighbors in P.

Firstly, differences, dj between p'i and its s nearest neighbors from P are calculated, then multiplied by a random number rj that lies between range of 0 and 1 (where j=1,2,…s) and finally s synthetic minority examples are generated between p'i and its nearest neighbors. Same procedure is done for each p'i and s × dnum synthetic instances are generated.

$$syn_j = p'_i + r_j \times d_j \text{ , where } j=1,2,...,s$$

Thus, generation of synthetic data along line segment between minority borderline examples and its nearest neighbors of the same class improves borderline examples. Borderline-SMOTE also generates examples synthetically from its nearest (negative) neighbor present in N. The difference is then multiplied by a random number that lies between range of 0 and 0.5, thus new generated samples lies closer to minority class.

III. HYBRID SAMPLING TECHNIQUES

The techniques that uses both oversampling as well as undersampling approaches to improve the imbalance ratio (IR) before classification process comes under the Hybrid Sampling approaches. The combination consists synthetic minority oversampling technique (SMOTE)along with its variants with tomek link used as undersampling technique. The techniques in the field of hybrid techniques stated below provide the emphasis on the removal of noisy examples and borderline examples such that accuracy and performance measures of the classifiers can be improved. It is done by using the concept of tomek link as the undersampling approach on the oversampled datasets.

A. SMOTE + TOMEK LINK

SMOTE +TL [15] was proposed by GE Batista et al. in 2004. In CIP, because of skewed data distribution, interpolating minority examples to oversample the minority class leads to introducing artificial minority class examples too deeply in the majority class space. Under such situation, a classifier can lead to problem of overfitting [16]. In order to obtain correct and better results,

applying Tomek link undersampling technique to the over-sampled training set works as a data cleaning approach.

This technique is illustrated as firstly, the original data set is over-sampled with Smote then Tomek links are identified. Finally, Tomek links are removed producing a balanced data set. The Smote + Tomek link technique is used to improve classification performance. It is used firstly for annotation of the proteins in field of Bioinformatics [15].

B. TOMEK LINK + CLUSTER SMOTE

Oversampling includes interpolating minority examples that leads to introducing synthetic minority class examples too deeply in the majority class space. Thus, it leads to the problem of overfitting. In order to obtain correct and better results, applying Tomek link as undersampling technique works as a data cleaning approach. Thus, removing only the majority class examples that form Tomek links are removed. After, which the Cluster SMOTE approach is applied to that reduced dataset.

This technique works in the following manner: Firstly, identify Tomek links in the original dataset and then remove them by using Tomek links as an undersampling approach and producing a dataset without noisy examples then dataset is over-sampled with Cluster-SMOTE producing well-defined class clusters.

The Tomek link + C-SMOTE technique improves the accuracy and performance results of classifiers than using C-SMOTE only for oversampling the data and reduce biasness towards majority class[18].

C. SAFE-LEVEL SMOTE + TOMEK LINK

Although over-sampling minority class examples can balance class distributions, but the presence noisy and borderline examples will provide hindrance to the accuracy of classifier. Therefore, applying Tomek links to the over-sampled training set as a data cleaning method provides better results as compared to only oversampled datasets. Thus, removing only the majority class examples that form Tomek links after oversampling the minority class with Safe level Smote is more efficient and predicts the rare class examples with more accuracy. The application of this technique is illustrated as:

Firstly, the original data set is over-sampled using SL- SMOTE and then Tomek links are identified and removed that produce a balanced data set with better performance measures results of classifier.

The SL- SMOTE + Tomek links technique improves the classification of examples in imbalanced datasets producing better accuracy and performance metrics than using SL- SMOTE only for oversampling the data and reduce the problem with imbalanced data sets[18].

D. BORDERLINE SMOTE + TOMEK LINK

Borderline Smote [17] method over-samples only the borderline samples of minority class. It helps in achieving better prediction values by learning the borderline of each class in dataset as exactly as possible in training process. Thus, Borderline-SMOTE is different from other approaches as it does not oversample all minority examples or a random subset of these minority class data. The idea of proposed technique is to improve Borderline SMOTE by removing the noisy examples and those borderline examples that belong to the majority class, with the help of identification of the Tomek links present in the oversampled data sets. This technique works as follows: Firstly, the original data set is over-sampled using Borderline SMOTE then Tomek links are identified and removed producing a balanced dataset.

The Borderline SMOTE + Tomek link technique improves the classification of examples in imbalanced datasets producing better accuracy and performance metrics than using Borderline SMOTE alone for oversampling data and reduce effects of CIP[18].

IV. CONCLUSION AND FUTURE SCOPE

In this paper, Class Imbalance Problem (CIP) is addressed at data level approaches. The oversampling techniques – Synthetic Minority Oversampling Technique (SMOTE), Cluster-SMOTE, Borderline SMOTE, Safe-level SMOTE, undersampling technique - Tomek link (TL) and hybrid sampling technique – SMOTE+TL, Tomek Link + Cluster SMOTE (TL+C-SMOTE),

Safe-Level SMOTE + Tomek Link (SL-SMOTE+TL) and Borderline SMOTE + Tomek Link (BorSMOTE+TL) is focused. The performance of these techniques are then compared using five publicly available datasets on UCI machine learning repository using Decision tree and K-Nearest Neighbor classifiers.

The classifier performance measures used for analyzing techniques are Sensitivity/Recall/TPR, Specificity/TNR, Precision, FPR, Overall Accuracy, F- Measure, G-mean and AUC (Area under ROC curve). Based on the results, it is observed that proposed hybrid sampling techniques- TL+C-SMOTE, SL- SMOTE+TL and BorSMOTE+TL outperforms their respective oversampling techniques – Cluster-SMOTE, Safe-level SMOTE and Borderline SMOTE.

REFERENCES

[1] Anjana Gosain and Saanchi Sardana, "Handling class imbalance problem using oversampling techniques: A review," International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2017, IEEE, pp. 79- 85, 2017.

[2] Guo Haixiang et al., "Learning from class-imbalanced data: Review of methods and applications", Expert Systems with Applications, Vol.73, pp. 220-239, 2017.

[3] Jiawei Han, Micheline Kamber and Jian Pei, Data Mining Concepts and Techniques, 3rd edition. Morgan Kaufmann Publishers,2011.

[4] Mikel Galar, Alberto Fern´andez, Edurne Barrenechea, Humberto Bustince, and Francisco Herrera,"A Review on Ensembles for the Class Imbalance Problem: Bagging, Boosting, and Hybrid-Based Approaches", IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, Vol. 42, No.-4, pp. 463-484, 2012.

[5] Mr.Rushi Longadge, Ms. Snehlata S. Dongre and Dr. Latesh Malik, "Class Imbalance Problem in Data Mining", Review International Journal of Computer Science and Network (IJCSN), Vol. 2, No. 1, February 2013.

[6] Prabhjot Kaur and Anjana Gosain, "Comparing the Behavior of Oversampling and Undersampling Approach of Class Imbalance Learning by Combining Class Imbalance Problem with Noise", In ICT Based Innovations, Springer, Singapore, pp. 23-30, 2018.

[7] Anjana Gosain and Deepika Singh, "Analysis of Sampling Based Classification Techniques to Overcome Class Imbalancing", 3rd International Conference on Computing for Sustainable Global Development Proceedings of the 10th INDIACom; INDIACom-2016, IEEE, pp. 2637-2643, 2016.

[8] N. V. Chawla et al., "SMOTE:Synthetic Minority Over Sampling Technique", Journal of Artificial Intelligence Research, Vol. 16, pp 321357, 2002.

[9] Paulraj Ponniah, Data warehousing fundamentals for IT professionals. John Wiley & Sons, 2011.

[10] N. V. Chawla, ''Data mining for imbalanced datasets: An overview,'' Data Mining and Knowledge Discovery Handbook. Springer, pp. 853– 867, 2005.

[11] David A Cieslak, Nitesh V Chawla and Aaron Striegel, "Combating Imbalance in Network Intrusion Datasets", [C]//GrC. IEEE,pp. 732-737, 2006.

[12] Yang Yong, "The research of imbalanced data set of sample sampling method based on K-means cluster and genetic algorithm", Energy Procedia 17, pp. 164- 170, 2012.

[13] C. Bunkhumpornpat, K. Sinapiromsaran, and C. Lursinsap, "Safe- LevelSMOTE: Safe Level- Synthetic MI Over-Sampling Technique for handling the Class Imbalance Problem", PADD2009, Springer LNAI, Vol. 5476, pp. 475-482, 2009.

[14] H. Han, W. Wang, and B. Mao, "Borderline-SMOTE: A new oversampling method in Imbalanced Data-sets Learning", In. ICIC 2005 LNCS, Springer, Heidelberg, Vol. 3644, pp. 878-887, 2005.

[15] GE Batista, RC Prati, MC Monard, "A study of the behavior of several methods for balancing machine learning training data",ACM SIGKDD explorations newsletter, Vol 6, No. 1, pp 20-29, 2004.

[16] Ajinkya More, "Survey of resampling techniques for improving classification performance in unbalanced datasets", arXiv preprint arXiv:1608.06048, 2016.

[17] N. V. Chawla, A. Lazarevic, L. O. Hall and K. W. Bowyer,"SMOTBoost: Improving pridiction of the minority class in boosting", In proc. Knowledge Discovery databases, pp. 107-119, 2003.

[18] Anjana Gosain, Arushi Gupta, and Deepika Singh,"Hybrid Data-Level Techniques for Class Imbalance Problem", International Conference on Innovative Computing and Communication, Springer, Vol 1, pp 1131-1141, 2020.

Windowing and Visual Simulation of its Protocols

Kanika Joshi [1], Ashish Sharma [2], Devender Banga [3]

[1] *Department of Information Technology*
Dr. Akhilesh Das Gupta Institute of
Technology & Management, New Delhi, 110053, India
[1]kanika1101999@gmail.com
[2]sharmaash989@yahoo.com
[3]*devender.banga@adgitmdelhi.ac.in*

***Abstract*— This paper specifically studies tools for network and protocol simulation. The project is aimed towards protocols of TCP that are used in data packet transfer from data layer to network layer. Windowing or Sliding Window is one of the major protocols that takes over the topic of transmission, where there is a sequential and reliable transmission of multiple data in form of packets and acknowledgment. The transfer of packets containing data is done with a sequence number and acknowledgement number which will provide assistance to both the windows it keeping the count of the packets sent and received.**

The loss in transmission is a major condition that affects the calculation of delay in data transmission. In addition to loss, it could be caused on either window. To mention, Acknowledgment from the receiver window plays a major role in the secure transmission of the data packet as it ensures the sender, the status of the packet transmitted.

Due to loss, delay or other factors there are various protocols that deal with such circumstance. To list we have to Stop & Wait, Go-Back N ARQ (Automatic repeat Request), Selective Request ARQ. These protocols were created to deal with the glitches caused during data transmission.

Our project is all about simulation and study of the data packet transmission under various circumstances of delay, loss and acknowledgment. In our project we will create a simulator with features to study various protocols at one terminal with the application of RSA Algorithm as an encryption method and verification of data transmitted using checksum.

***Keywords*— Transmission, Window, Packets, Acknowledgment, Delay, Loss**

I. INTRODUCTION

For data to be delivered correctly at its destination, is a task in an unreliable channel, such as in the data link layer (OSI layer 2) and in TCP [9]. One of the most important aspect of using windowing in TCP that it provides a secure, reliable & sequential delivery of frames over windows [10]. In this protocol each frame has been sent with a 3-bit sequence number to the receiver's window, to which a 3-bit acknowledgment number is sent back to the sender which signifies it's delivery at the intended destination. The performance of transmission is affected by various aspects of communication channel. While transmission data could be lost, duplicated, or could reorder the messages [7]. For such problems a well-studied, well-defined protocol comes as a satisfactory solution, i.e. Sliding Window Protocol. The flow of the data packets is satisfactorily handled by Sliding Window Protocol [6].

1.1 Stop & Wait Protocol:

This protocol is implemented when a packet has been released with sequence number and data frame and is waiting for its acknowledgement from the receiver window. The next packet will wait until the acknowledgment of the previous is received [8].

1.2 Sliding Window Protocol:

This protocol takes the charge of secure transmission of packets transmitted into the communication channel of Data Link layer (OSI Layer 2) as well as of TCP [9]. It was presented to overcome what lacked in Stop & Wait Protocol.

1.2.1 Go- Back N ARQ (Automatic Repeat Request):

This protocol was introduced to overcome the problem of data corruption or lost data. [11] In this protocol all the subsequent frames have to be sent again from the one which was successfully transmitted in order to complete the transmission. It was also observed that this protocol has a high rate of error and wastage of bandwidth [11].

1.2.2 Selective Request ARQ:

To overcome the problems faced in Go-Back N ARQ, this protocol was introduced in which the frame which is corrupted or lost is sent again only as it solves the problem of wastage of bandwidth and had low error rates.

II. METHODOLOGY

The protocol is based on window either it is sender window or the receiver window. In which, window size plays a major role in deciding how many bits of data can be transmitted in the channel. On the basis of window size as assigned 0 to 2^n-1 [10], we have:

A. 1-bit Windowing

In this, the window size is of 1. This is applicable for Stop and Wait Protocol, which means sending 1 data packet and wait for the next until its acknowledgment is received from the receiver's window.

B. 2-bit Windowing

In this, the window size is of 4. In this 4 data packets can be sent simultaneously without waiting for its acknowledgment.

C. 3-bit Windowing

In this, the window size is of 8. In this 8 data packets can be sent simultaneously without waiting for its acknowledgment.

III. IMPLEMENTATION

Our project is a learner tool in which we aimed to portray:

3.1 Ideal Case:

It will have all the ideal criteria of transmission i.e. packet sent securely with its acknowledgment received back without delay.

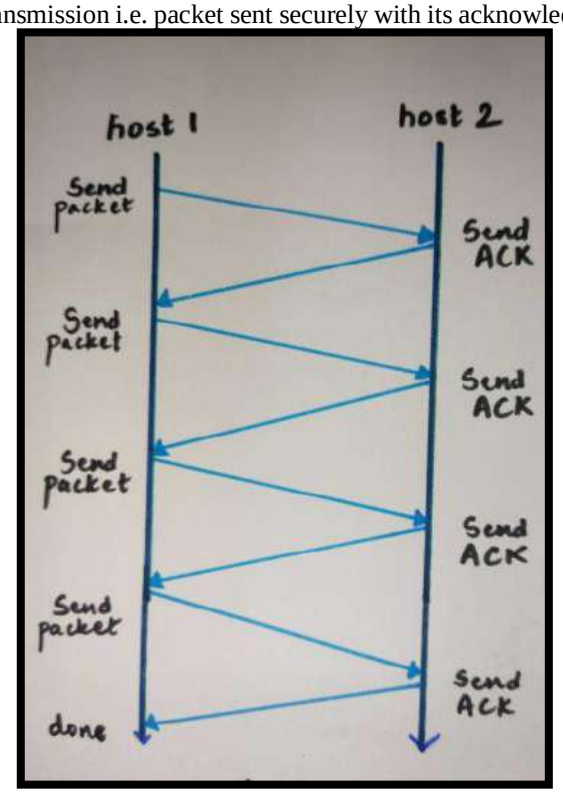


Fig. 1 Case (A): 1-bit Transmission

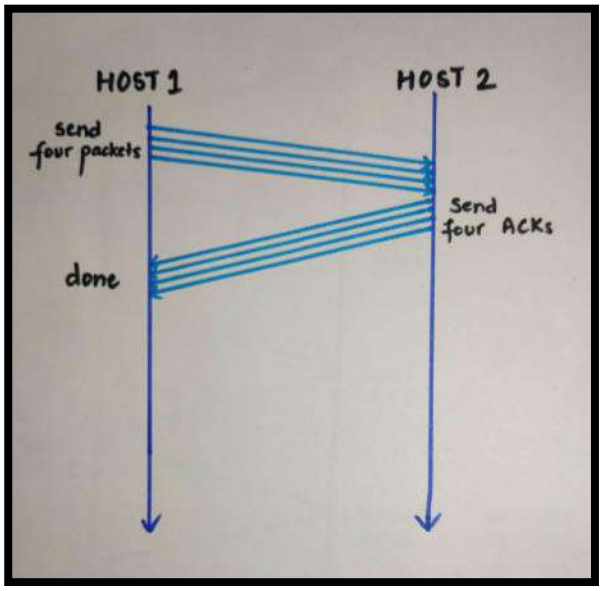


Fig. 2 Case (B): Multiple Packet & Acknowledgement Transmission

3.2 Packet Loss:

It will have condition where a packet will be lost while transmission from sender to receiver. For which the sender waits for its acknowledgment and retransmits the packet after time out.

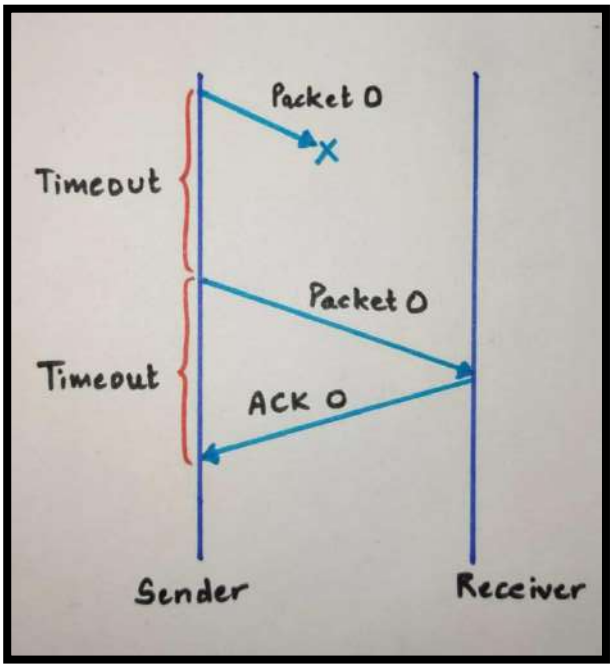


Fig. 3 Packet loss while Transmission

3.3 Acknowledgement Loss:

It will have condition where a packet is successfully transmitted from sender to receiver but no acknowledgment for the same is received. In such case the packet will be retransmitted after timeout.

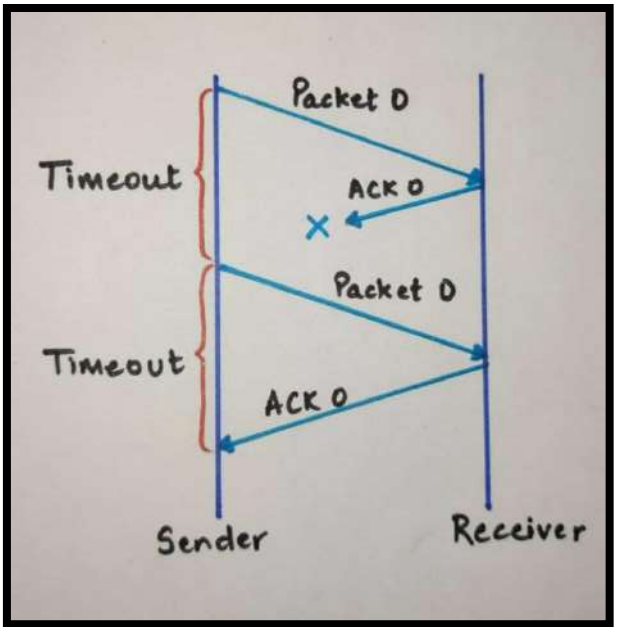


Fig. 4 Acknowledgement Loss while Transmission

DATA SECURITY – RSA ENCRYPTION ALGORITHM

RSA (Rivest–Shamir–Adleman) algorithm, a public key encryption a type of asymmetric cryptography algorithm, is based on prime numbers. This algorithm works on two key method, i.e. Public key and Private Key. Public key is the key known by all the hosts, while private key is the key which helps the intended host to decrypt the encrypted data received from the sender host to enhance the security of the data packet transmitted by factorization of the variables [2].

For the Public and private keys we will use JAVA editor that will be used in the implementation of the same.

Steps to Implement RSA Algorithm:

Encryption

Step 1: Calculating the number of messages *N* to be encrypted

$$N = a*b \quad (1)$$

where, *a* and *b* are the random numbers used to calculate the modulus for the function of encryption.

Step 2: Calculating the function of encryption of the messages using the prime numbers

$$f(N) = (a-1)(b-1) \quad (2)$$

Step 3: Calculating the exponential *e* used for the calculation of Cipher in encryption selection of the same should be done such that $1<e<f(N)$ and $GCD(e, f(N))=1$.

Step 4: Public Key = (N, e)

Step 5: Calculating the Cipher text *C* used for encryption

$$C = P^{e} \bmod N \quad (3)$$

where, P is the plain text/ message text.

Decryption

For the receiver to read the encrypted message, decryption of the cipher is needed which is done with the help of private key available with the receiver host.

Step 6: Determining the private key K at the receiver

$$K = mod\ (f(N))/e \quad (4)$$

Step 7: Private Key = (N, K)

Step 8: Decryption of the cipher text received from the sender host

$$P = C^{K} \bmod N \quad (5)$$

Here after decrypting, the cipher text it is turned back to its original form.

IV. DATA INTEGRITY – CHECKSUM

Error detection of the data transmitted is an important parameter that defines the integrity of the transmission of the data. Checksum is a way of verification which is usually done using the comparing the original set of message/data with the received set of the message/data to verify both are same.

In this project, we applied the checksum on both the sender' and receiver' end in which a segment is applied on the data frame using 1's compliment to get the sum.

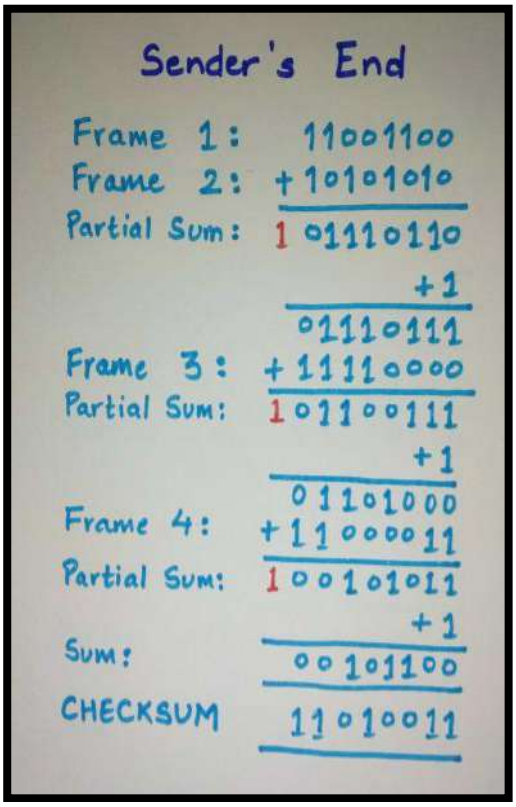


Fig. 5 Checksum at the Sender Window

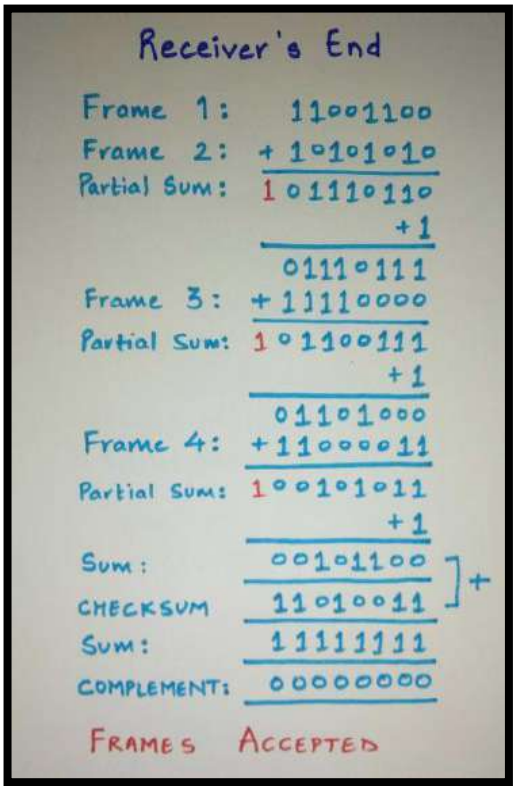


Fig. 6 Checksum at the Receiver Window

V. RESULT

The simulator designed will have sending and receiving window. For packet transmission, we have buttons such as START, STOP, PAUSE, HELP, and CHANGE SETTINGS. Using the settings button, we could change our settings of sequence number, losses on the way of transmission.

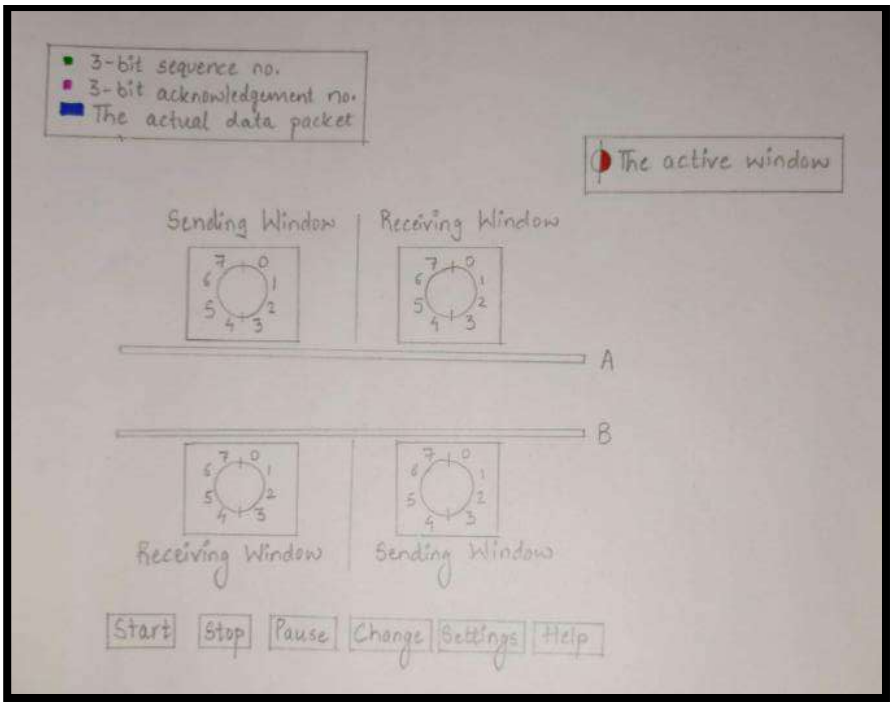


Fig. 7 Initial Stage of Stimulator before Transmission

Packet transmission is been shown the in the window, for which settings could be changed according to the requirement. Each packet will contain Sequence number, Acknowledgment number, and data packet [7].

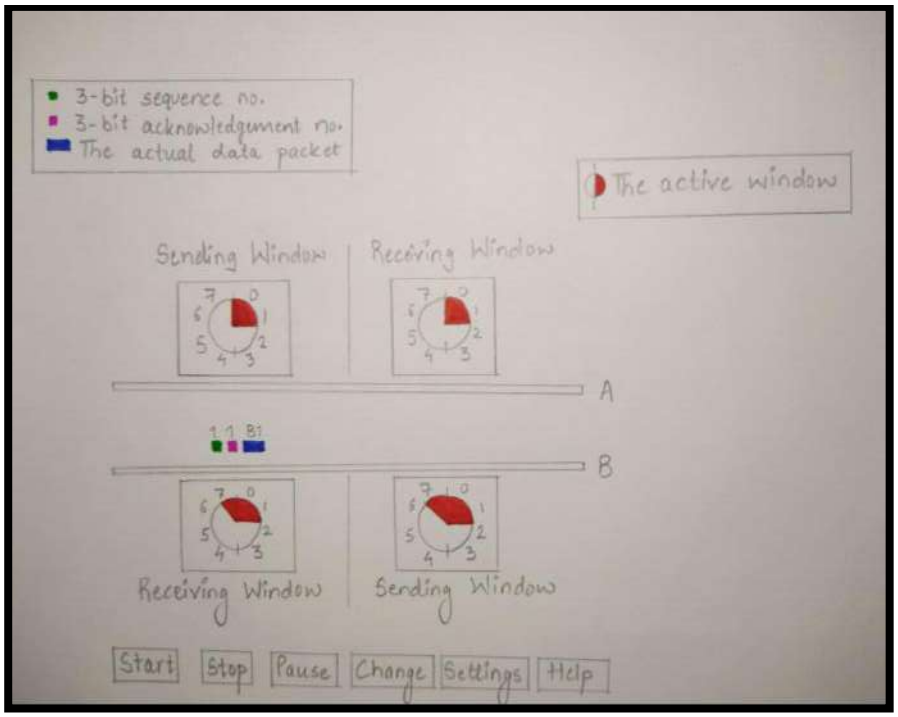


Fig. 8 Initial Stage of Stimulator before Transmission

For the security, RSA Encryption algorithm was successfully applied by generating the public and private keys. And for the integrity of the data in transmitted frames we applied checksum successfully.

VI. Conclusion

In this project we have achieved to implement transmission of 1-bit, 2-bit and 3-bit frames under Sliding Window Protocol. In this paper we worked on the data security by generating public and private keys with the application of RSA Algorithm and for data integrity checksum was applied on the frames on both receiver' and sender' side.

Acknowledgment

It gives us massive delight to manifest our appreciation to the people who have hand out in successful completion of the project. We owe a debt of gratitude to Mr. Devender Banga, Assistant Professor, Dr. Akhilesh Das Gupta Institute of Technology & Management, who provided us virtuous resources for this project. We are highly indebted to our project guide.

References

[1] Abhishek Guru and Asha Ambhaikar, "Development of RSA Encryption Algorithm for Secure Data Transmission," Research Journal of Computer and Information Technology Sciences, vol. 8(1), pp. 2320–6527, June 2020.
[2] Taleb Samad Obaid, "Study A Public Key in RSA Algorithm," European Journal of Engineering Research & Science, vol. 5(4), pp. 395–398, April 2020.
[3] Varun Shukla, Atul Chaturvedi, Neelam Srivastava, "A Secure Stop and Wait Communication Protocol for Disturbed Networks," Springer Science plus Business Media, LLC, part of Springer Nature 2019, vol. 2(1), September 2019.
[4] Hoang D. Le, Vuong Mai, Chuyen T. Nguyen, Anh T. Phem, "Design and Analysis of Sliding Window ARQ Protocols With Rate Adaptation for Burst Transmission Over FSO Turbulence Channels Journal of Optical Communication and Networking, vol. 11(5), pp. 151–163, May 2019.
[5] Tsado Victor Mamman, Sadiq Umar (PhD), Ashigwuike Evans (PhD), Vitalis C. Anye, "A Review Of Selective Repeat Arq Scheme For Channel Error Correction In Point To Multipoint Service," International Journal of Scientific & Engineering Research, vol. 10(3), pp. 2229–5518, March 2019.
[6] Ayushi Chaudhary, "Exploring of Sliding Window Visualization System to Understand Flow and Error Control Mechanism of Data Link Layer," International Journal of Soft Computing and Engineering, vol. 5(4), pp.107–111, September 2015.
[7] Abhilasha Sahu, Anubhuti Pandey, Gourav Shukla, R.Chetan, Kiran Chhabra, "Simulation of Sliding Window Protocol," International Journal of Computer Science and Network, vol. 5(4), pp. 198–200, May 2014.
[8] Kapil Kumar Bansal, Rakesh Kumar Yadav, "Analysis of Sliding Window Protocol for Connected Node," International Journal of Computer Science and Network, vol. 2(5), pp. 292–294, November 2012.
[9] Dmitri Chkliaev, Jozef Hooman, and Erik de Vink, "Verification and Improvement of the Sliding Window Protocol," unpublished.
[10] Samual Sam, "Sliding Window Protocol," Internet: www.tutorialspoint.com/sliding-window-protocol, February 20, 2019.
[11] JavaTpoint, " Sliding Window Protocol," Internet: javatpoint.com/sliding-window-protocol, July 30, 2018.

Fuzzy Clustering and Kernel-Based Type-2 Fuzzy Clustering: A comparative Study

Saijal Gupta[1], Monica Batra[2], Gunjan Chugh[3]

Assistant Professor, Dept. of IT, ADGITM

GGSIP University, New Delhi, India

[1]saijalgupta92@gmail.com

[2]monica.batra88@gmail.com

[3]chugh.gunjan8917@gmail.com

***Abstract* — Fuzzy or soft clustering is used in the clustering of data, when a point having different degrees of membership can belong to two or more clusters. This technique is applied in some common areas like detection of crime, medical images, pattern recognition etc. In this paper, we present the experimental analysis of all the major fuzzy clustering techniques and review their performances. The algorithms are: FCM, FCM- σ, T2FCM, T2FCM- σ, KFCM and KT2FCM. Anew algorithm has been proposed named as KT2FCM- σ. A standard dataset is chosen to have better analysis. Kernel based Type-2 algorithms uses a mapping function and hence gives better performance. This paper will help researchers to pick an appropriate algorithm who are dealing with fuzzy clustering.**

***Keywords*— Fuzzy Clustering, FCM, T2FCM, KT2FCM, KT2FCM- σ, RBF Kernel Based fuzzy c-means**

I. INTRODUCTION

An unsupervised learning method to group a set of abstract or physical objects into classes of identical objects is called clustering. A cluster has the property that similar data objects are collected in the same cluster and differing data objects in another cluster. A cluster of data objects is collectively considered as one group and thus a form of data compression. The clustering aims on the set of unlabeled data and determines the intrinsic grouping [3,5].

Clustering of data is the basis of many classification algorithms. The main aim of clustering is to recognize the data set showing similar properties from a large dataset and form their groups or clusters [7]. Clustering is categorized into two forms: Hard clustering and Soft or Fuzzy Clustering. In hard clustering, each element can reside into exactly one cluster. In the case of soft clustering, clustering is done using fuzzy datasets. Hence, each element may belong to more than one cluster with distinct degrees of membership. Thus, the data is correlated to a membership value. Mostly, hard clustering is less natural than soft clustering [1]. The objects lying on the boundaries of various groups or clusters are not compelled to move to one of the classes, but rather are given partial membership by assigning a membership degree with value ranging from 0 to 1. On the contrary, in hard clustering techniques, data is exclusively grouped such that it can only belong to one cluster and hence cannot be included in another cluster[6]. The most famous fuzzy clustering technique is Fuzzy C Means (FCM) and similarly an important technique of hard clustering is K-means. These fuzzy sets can be categorized as either Type-1, Type-II or Intuitionistic.

The FCM algorithm developed by Bezdek, and its variants like FCM-$\boldsymbol{\sigma}$, PFCM, PCM which were derived from FCM have been useful in pattern recognition, image analysis. Although they give best results for overlapped data set and with noise free data but they fail to recognize the actual data and the outliers or noise. In the presence of an outlier, the centroid is attracted towards them. Similarly, PCM and PFCM were also not accurate in finding the optimal clusters when noise or outlier is present.

KFCM (Kernel Fuzzy c-means Clustering), was obtained by replacing the original Euclidean metric in FCM with a new kernel-induced parameter, still keeping the clustered prototypes in the data space to reformulate and interpret the clustering results in the original data space. KFCM has the tendency to tolerate unequal sized clusters and is robust to outliers and noise.

KFCM-$\boldsymbol{\sigma}$ shows improvement over KFCM on real data sets. Kernel-based clustering techniques are useful to partition non-linear distributed data, but only upto quadratic functions. The algorithms based on Type-2 fuzzy data sets are T2FCM and KT2FCM. Their performance is better than earlier proposed algorithms, but T2FCM fails for non-spherical and complex input patterns. KT2FCM, by introducing kernel and tangent functions, overcomes this problem and performs better over a noisy data. The issue with these algorithms is that they are inefficient in finding the accurate clusters. To improve the accuracy, a new algorithm KT2FCM-$\boldsymbol{\sigma}$ is introduced into the mapped feature space by integrating a new distance measure.

In this paper, we have described these algorithms in detail and a comparison in their performances is showed. The paper has the following organization: In section 2, we made a discussion of all the fuzzy clustering techniques and the new proposed technique of KT2FCM-$\boldsymbol{\sigma}$ is discussed in section 3. In section 4, evaluation is done on a standard dataset and its description is given for experimental analysis in the form of figures. In section 5, the results are represented and centroids of all the algorithms are tabulated. At last, conclusions and future scope is given in section 6.

II. LITERATURE SURVEY

In this section, various fuzzy clustering algorithms have been discussed. The input data-set denoted by 'X', given by X={x1, x2, x3,.. xn}, specifies n points in 2-dimensional space divided into 'c' clusters. Centroids of the clusters are denoted by vk and k is used to specify the cluster. Therefore, the value of k ranges from [1,c].

Fuzzy C-Means(FCM) [14]

The clustering technique developed by Dunn, was FCM, which was enhanced by Bezdek. The assumption in this algorithm is that the number of clusters, 'c', is predetermined for the given dataset and thus the objective function is:

$$J = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ik}\ \|x_i - v_k\|^2 \qquad (1)$$

Where, v_k is the centre vector, u_{ik} is the membership degree. The norm, $\|x_i - v_k\|$ is denoted by 'd_{ik}'.
The degree of membership for (1) is obtained as [4]:

$$u_{ik} = \frac{1}{\sum_{k=1}^{C}\left(\frac{d_{ki}}{d_{ji}}\right)^{\frac{2}{m-1}}} \qquad (2)$$

Where, m is the fuzziness coefficient and the center vector v_k is calculated as follows:

$$c_k = \frac{\sum_{i=1}^{N} u_{ik}^m . x_i}{\sum_{i=1}^{N} u_{ik}^m} \qquad (3)$$

But, FCM is inefficient in detecting the accurate centroids when noise and outliers are present.

Fuzzy C-Means with new distance metric(FCM-$\boldsymbol{\sigma}$)[8]

D.M Tsai and C.C Lin[10] proposed the distance metric used in FCM and changed to a new metric, d'^2_{ki} which is defined as:

$$d'^2_{ki} = \frac{x_i - v_k}{\sigma_k}. \qquad (4)$$

Where σ_k is the weighted mean distance of cluster center v_k and is defined as:

$$\sigma_k = \sqrt{\frac{\sum_{i=1}^{n} u_{ik}^m \|x_i - v_k\|^2}{\sum_{i=1}^{n} u_{ik}^m}} \qquad (5)$$

and the minimized objective function J is:

$$J' = \sum_{k=1}^{C}\sum_{i=1}^{n} u_{ik}^m \frac{\|x_i - v_k\|^2}{\sigma_i} \qquad (6)$$

Membership function is defined as:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{d'_{ki}}{d'_{ji}}\right)^{\frac{2}{m-1}}} \quad \forall i,k \qquad (7)$$

Here, the integer k belongs to the range[1,c] and i is also an integer in the range of [1,n].

The cluster equation thus becomes:

$$v_k = \frac{\sum_{i=1}^{n}(u_{ik}^m\, x_k)}{\sum_{i=1}^{n}(u_{ik}^m)} \qquad \forall i \qquad (8)$$

Type-2 Fuzzy C-Means(T2FCM) [9]

The algorithm was proposed by Hwang and Rhee [15]. The Type-2 membership is evaluated as [11]:

$$a_{ik} = u_{ik} - \frac{1 - u_{ik}}{2} \qquad (9)$$

where a_{ij} is the Type-2 membership used in the centroid computation. All the other computations are same as that of the conventional FCM. The equation used for the calculating the cluster center 'v_j' [8]:

$$v_k = \frac{\sum_{i=1}^{n}(a_{ik}^m\, x_i)}{\sum_{i=1}^{n}(a_{ik}^m)} \qquad (10)$$

Centroids obtained from Type-2 sets are more accurate than the Type-1 sets.

Type-2 Fuzzy C-Means with new distance metric (T2FCM- $\boldsymbol{\sigma}$*) [9]*

The distance metric in the case of T2FCM was changed to a new metric, $d'_{ki}{}^2$ which is defined as:

$$d'^2_{ki} = \frac{x_i - v_k}{\sigma_k} \qquad (11)$$

The equation used for updating the cluster centers 'v_k' [8] is as follows:

$$v_k = \frac{\sum_{i=1}^{n}(a_{ik}^m\, x_i)}{\sum_{i=1}^{n}(a_{ik}^m)} \qquad (12)$$

All other equations are calculated as FCM-σ. It performs better over noisy data but fails in non-spherical structures.

Kernal Fuzzy C-Means(KFCM)[12]

KFCM was proposed to obtain accurate clusters using the RBF functions, defined as:

$$K(x, y) = \exp\left(-\sum_i \frac{|x_i^a - y_i^a|^b}{\sigma^2}\right) \quad (0 < b \leq 2) \qquad (13)$$

The objective function is:

$$J = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^m \|\emptyset(x_k) - \emptyset(v_i)\|^2 \qquad (14)$$

Where

$\|\emptyset(x_k) - \emptyset(v_i)\|^2 = K(x_k, x_k) + K(v_i, v_i) - 2K(x_k, v_i)$

where K(x,y) is the kernel function. Also, K(x,x)=1, thus minimizing the equation as:

$$J = 2\sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^m (1 - K(x_k, v_i)) \qquad (15)$$

The cluster center v_i is obtained as:

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^m K(x_k, v_i) x_k}{\sum_{k=1}^{n} u_{ik}^m K(x_k, v_i)} \qquad (16)$$

The membership matrix is as follows:

$$u_{ik} = \frac{\sum_{k=1}^{n}(1 - K(x_k, v_i))^{\frac{1}{(m-1)}}}{(1 - K(x_k, v_i))^{\frac{1}{(m-1)}}} \qquad (17)$$

Kernelized Type-2 Fuzzy C-Means(KT2FCM)[10]

KT2FCM is proposed by exploiting the Radial Basis Function Kernel (RBF) used in KFCM with the Type-2 fuzzy sets.

The objective function is thus minimized to [13]:

$$J = 2\sum_{i=1}^{c}\sum_{k=1}^{n} a_{ik}^{m}\|\emptyset(x_k) - \emptyset(v_i)\|^2 \qquad (18)$$

Where a_{ik} is the Type-2 membership.

Also, a_{ik} is defined as in the case of T2FCM as:

$$a_{ik} = u_{ik} - \frac{1 - u_{ik}}{2} \qquad (19)$$

The cluster center is computed as:

$$v_i = \frac{\sum_{k=1}^{n} a_{ik}^{m} K(x_k, v_i) x_k}{\sum_{k=1}^{n} a_{ik}^{m} K(x_k, v_i)} \qquad (20)$$

The centroids obtained were more accurate by the introduction of kernel in the Type-2 fuzzy sets.

III. THE PROPOSED TECHNIQUE

Kernelized Type-2 Fuzzy C-Means with new distance metric (KT2FCM- $\boldsymbol{\sigma}$*)*

The distance metric used in conventional FCM and T2FCM is used with KT2FCM to obtain more accurate results for the centroid. The metric is defined as:

$$d'^{2}_{ki} = \frac{x_i - v_k}{\sigma_k} \qquad (21)$$

Thus, the membership function is defined as:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c}\left(\frac{d'_{ki}}{d'_{ji}}\right)^{\frac{2}{m-1}}} \quad \forall i, k \qquad (22)$$

Where integer k is in the range [1,c] and I is in the range of [1,n].

The objective function after minimizing is as follows:

$$J = 2\sum_{i=1}^{c}\sum_{k=1}^{n} a_{ik}^{m}(1 - K(x_k, v_i)) \qquad (23)$$

IV. EVALUATION ON DATASET

In this section of evaluation, a standard dataset has been taken for the experimental analysis and to show the performance comparison. It is represented in Fig. 1 consisting of 140 data points.

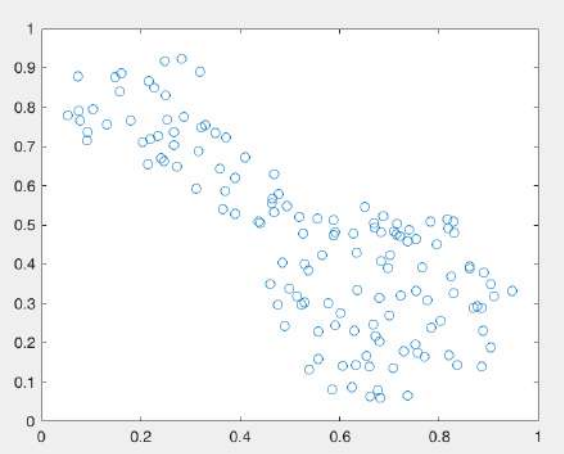
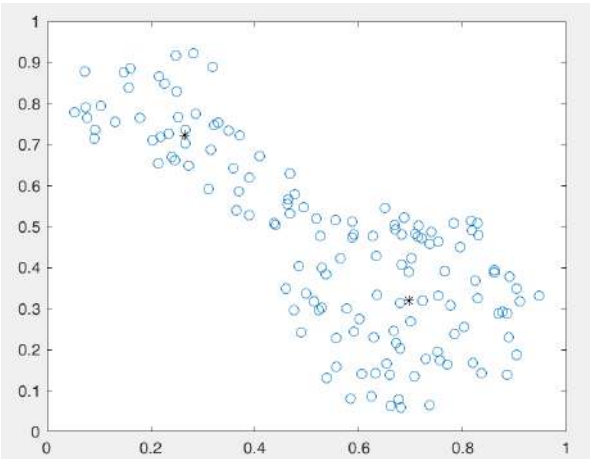
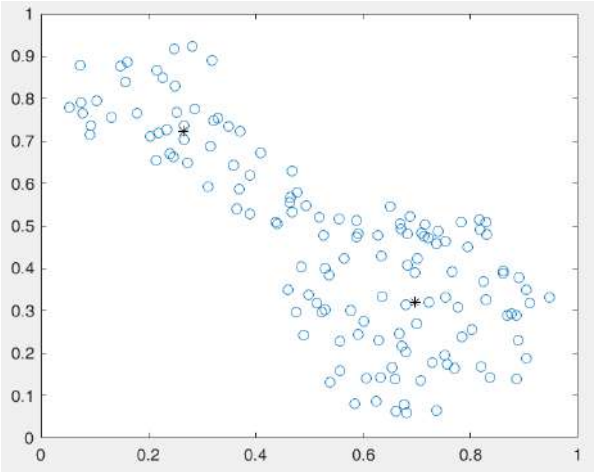


Fig. 1 Fcmdata Dataset

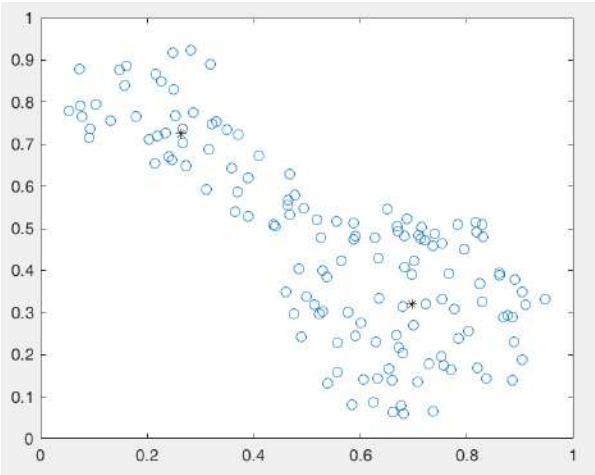
The centroids for all the algorithms are represented as:



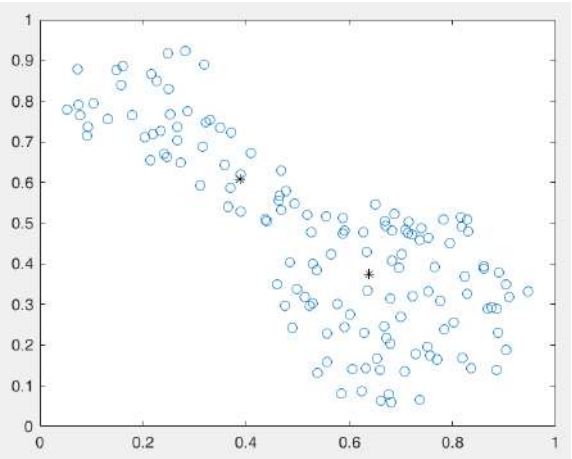
(a)



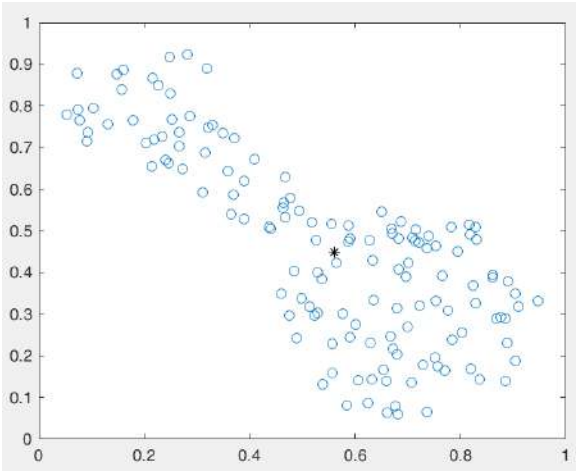
(b)



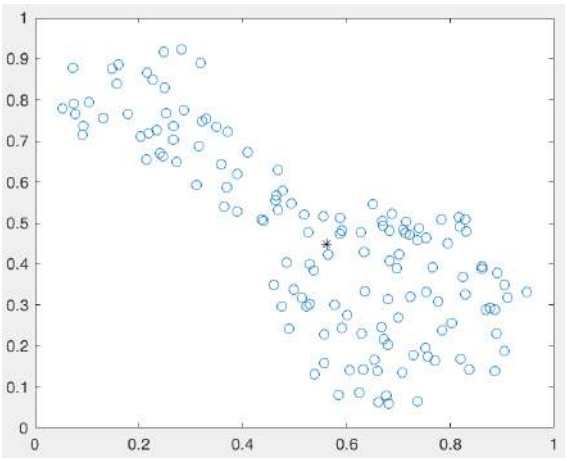
(c)



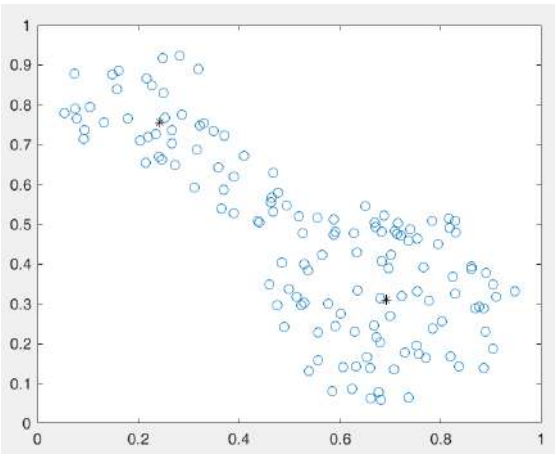
(d)



(e)



(f)



(g)

Fig. 2 (a) FCM; (b) FCM- σ ; (c) T2FCM; (d) T2FCM- σ
(e) KFCM; (f) KT2FCM; (g) KT2FCM- σ

V. Experimental Results and Discussions

The centroids for the algorithms discussed have been represented in Fig. 2. The input data points are represented with rounds and the centroids obtained are shown by '*'. The coordinates of the centroid obtained are tabulated. C1 and C2 are the centroids for cluster 1 and cluster 2, respectively. From Fig. 1 and Table 1, it is observed that the algorithm proposed (KT2FCM- σ) gives the most accurate results for centroids. The centroids obtained are closest to the actual centroid values. A large difference is seen in KFCM and KT2FCM due to the attraction of centroid to noise and outliers. So, it can be proved from table 1 that the proposed algorithm gives best results and is highly accurate.

TABLE 1. COORDINATES OF CENTROID ON THE DATASET

Algorithm Used	No. of Clusters	C1		C2	
FCM	2	0.265448719119249	0.722944206110164	0.696869369156904	0.320562511287800
FCM- σ	2	0.265446381618595	0.722946333226066	0.696868274923825	0.320563681198494
T2FCM	2	0.262825015485002	0.725437090331602	0.697165543280520	0.320012367562726
T2FCM- σ	2	0.387386091247875	0.608822820231453	0.638330253903576	0.375034022528131
KFCM	2	0.561137433459052	0.448466196042623	0.560553874170677	0.449019497309122
KT2FCM	2	0.561461037084402	0.448241372142590	0.561395702371298	0.448303692481036
KT2FCM- σ	**2**	**0.241754343894414**	**0.756621677146743**	**0.692471552982933**	**0.310413300353630**

VI. CONCLUSIONS AND FUTURE SCOPE

In this paper, we analyzed all the fuzzy clustering algorithms on a standard dataset and proposed a kernel-induced new metric for the Type-2 fuzzy c-means algorithm. It has been observed that FCM- σ improves the centroid than FCM. T2FCM and T2FCM- σ significantly improves the performance over FCM and FCM- σ. KFCM and KT2FCM are more robust to noise and thus improves centroid position. KT2FCM- σ gave best results in the calculation for centroids when compared to other algorithms.

Kernel-based clustering techniques are useful to partition non-linear distributed data, but only upto quadratic functions. So, as a future work it can be extended to higher polynomial function. Also, introducing the concept of hesitation degree will improve the performance.

REFERENCES

[1] Tan, Steinbach, Kumar (2004) Data Mining Cluster Analysis: Basic Concepts and Algorithms. [Online]. Available at: https://www-users.cs.umn.edu/~kumar/dmbook/dmslides/chap8_basic_cluster_analysis.pdf

[2] Han, J., Kamber, M. (2001) Data Mining: Concepts and Techniques, 2nd edn.: Morgan Kaufmann Publishers.

[3] Manish Verma, Mauly Srivastava, Neha Chack, Atul Kumar Diswar, Nidhi Gupta, "*A Comparative Study of Various Clustering Algorithms in Data Mining*", International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com, vol. 2, Issue 3, pp.1379-1384,May-Jun. 2012

[4] Yaikhom, Gagarine (2016) Yaikhom, Available at: http://yaikhom.com/2013/03/16/implementing-the-fuzzy-c-means-algorithm.html (Accessed: 3rd Oct 2019).

[5] Osama Abu "*A Comparison between data clustering algorithms*", The International Arab Journal of Information Technology, Vol.5, No: 3, pp: 320-325, 2008.

[6] Jyoti Bora, D., Dr. Gupta, A. (n.d.) '*A Comparative study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm'*, International Journal of Computer Trends and Technology (IJCTT) , 10(2), pp. 108-113.
[7] Gosain, A., Dahiya, S. (2016) *Performance Analysis of Various Fuzzy Clustering Algorithms*: A Review, Elsevier B.V.
[8] Chaira, Tamalika. "*A novel intuitionistic fuzzy C means clustering algorithm and its application to medical images*." Applied Soft Computing 11, no. 2 (2011): 1711-1717.

[9] Rhee, Frank Chung Hoon, and Cheul Hwang. "*A type-2 fuzzy C-means clustering algorithm*." In IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th, vol. 4, pp. 1926-1929. IEEE, 2001.

[10] Kaur, Prabhjot, I. M. S. Lamba, and A. Gosain. "*Kernelized type-2 fuzzy c-means clustering algorithm in segmentation of noisy medical images*." In Recent Advances in Intelligent Computational Systems (RAICS), 2011 IEEE, pp. 493-498. IEEE, 2011.

[11] Kaur, Prabhjot, A. K. Soni, and A. Gosain. "*Novel intutionistic fuzzy c means clustering for linearly and non-linearly separable data*." WSEAS Trans Comput 11 (2012): 65-76.

[12] Zhang, Dao-Qiang, and Song-Can Chen. "*Clustering incomplete data using kernel-based fuzzy c-means algorithm*." Neural Processing Letters 18.3 (2003): 155-162.

[13] Dave, Rajesh N. "*Robust fuzzy clustering algorithms*." In Fuzzy Systems, 1993., Second IEEE International Conference on, pp. 1281-1286. IEEE, 1993.

[14] J.C. Bezdek (1981), "*Pattern Recognition with Fuzzy Objective Function Algorithm*", Plenum, NY.

[15] *F.C.H. Rhee, C. Hwang, A Type-2 fuzzy c means clustering algorithm*, in: Proc. in Joint 9th IFSA World Congress and 20th NAFIPS International Conference 4, 2001, pp. 1926–1929.

Deep Learning Can be used to Detect Fake News? A Survey: for Online Social Platform

Monica Batra[1], Saijal Gupta[2], Gunjan Chugh[3]
Assistant Professor, Dept. of IT, ADGITM
GGSIP University, New Delhi, India
[1]monica.batra88@gmail.com
[2]saijalgupta92@gmail.com
[3]chugh.gunjan8917@gmail.com

***Abstract* - In a systematic literature review method, we recently, found the usage of social networking websites etc. On online social networks such as Facebook, Twitter, and LinkedIn which has become an inseparable part of our daily lives. It is considered as a easy platform for all users to share personal messages, pictures, and videos. However, while people enjoy social networks, many activities which are deceptive such as fake news or rumors can mislead users to believe it right. Our main target is mapping the state of art of fake news detection, defining fake news and finding the most useful deep learning technique for doing so. We concluded that the most used method for automatic fake news detection is not just one classical deep learning technique, but instead a amalgamation of classic techniques coordinated by a neural Therefore, related to the previous issues, we present a comprehensive survey of automated misinformation detection on (i) false information, (ii) rumors,spam, (iv) fake news, and (v) disinformation. We will provide a state-of-the-art review on MID where deep learning (DL) is mainly used to automatically process data and create patterns to make decisions not only to extract global features but also to achieve good results.**

***Keywords* - deeplearning, social network,fake news,rumour,machine learning**

I. INTRODUCTION

Different from the beginning of the internet, we produce more data and information than we are able to consume. Consequently, it's potential that some information or rumors square measure generated and unfold throughout the net,[1] leading different users to believe and propagate them, during a chain of unintentional (or not) lies. Such information will generate unreal thoughts and opinions, collective hysteria or different serious consequences. so as to avoid such things to happen, specially closed to political events like elections, researchers are finding out the data flow and generation on social medias within the last years, that specialize in subjects as opinion mining, users relationship, sentiment analysis, emotion unfold,Twitter, and linkden, people share their opinions, videos, and news on their various activities, while people enjoy social networks, many deceptive activities such as fake news, or rumors can mislead users into believing misinformation in this context as they have gained great popularity in the last three years[5].

1. The existing surveys covered a broad range of techniques used for MID. However, given the increasing popularity of using DL methods to detect misinformation, we believe our survey provides a timely review of the use of DL techniques. For example, we reviewed how different MID problems are covered under various DL techniques, which were not covered in existing surveys. We hope this survey can benefit researchers to deep insight between related techniques and these issues.

2. We present a state-of-the-art systematic review of the existing problems, solutions, and validation of MID in online social networks based on various DL techniques.
3. To identify the recent and future trends of MID research, we analyze the key strengths and limitations of the existing various techniques and describe the state-of-the-art DL as an emerging technique on massive social network data.
4. In the rest of the paper, first, we present the formal problem definition, types, impacts, and DL with the associated challenges. We then present the state-of-the-art DL techniques for MID.

II. MISINFORMATION

Misinformation is ―false information that is spread, regardless of intent to mislead Misinformation vs.-Disinformation: Get Informed on The Difference Amid the coronavirus pandemic, we are all desperate for information. Where did the virus return from? Is there a cure? however will we tend to keep staying safe? can life come back to to normal? within the case of COVID-19, ,[2] data is a literal life-saver—when it's true. Wrong data doesn't facilitate anyone and may even create things worse. And sort of a virus, wrong data will unfold, inflicting what's been known as Associate in Nursing infodemic. False and dishonest data unfold like ne'er before in 2020.[16]

Learn a lot of regarding however that affected our language in 2020 Word of the Year article. And currently quite ever, we tend to square measure seeing the unfold of 2 kinds of wrong information: info and misinformation. These 2 words, therefore usually used interchangeably, square measure just one letter apart. however, behind that one letter hides the vital distinction between these similar words: intent.

The spread of misinformation happens often in our everyday lives. We human beings—news flash—are not perfect [3]. We can all make mistakes. We all forget things. We mishear or misremember details. We tell our friends something we heard on TV or saw on social media that wasn't really true. If you are spreading around information that is wrong but you don't know it is wrong, then you are, well, technically, spreading misinformation. And when we say misinformation is an everyday thing, we mean it. For example, say a party starts at 8pm, but you forget or misread the invitation and tell your friends it starts at 9pm, you are supplying them with misinformation.

III. TYPES OF MISINFORMATION

Fig.1 Types of Fake news

IV. DEEP LEARNING

Deep learning is a subfield of machine learning that deals with algorithms induced by the structure and function of the brain, called artificial neural networks. Deep learning algorithms try to exploit the unknown structure in input distribution to find good representations, often, at many levels, high-level learning features are defined in terms of lower-level features. Deep learning is the key to performing complex tasks at a high level of sophistication. However, successfully constructing and deploying them proves to be a challenge for scientists and engineers across the data. [7] Although data training takes a little longer, testing can be done in a much shorter time. To accelerate the DL process, the DL Framework combines the implementation of modulated DL algorithms, optimization technology, distribution technology and support infrastructure

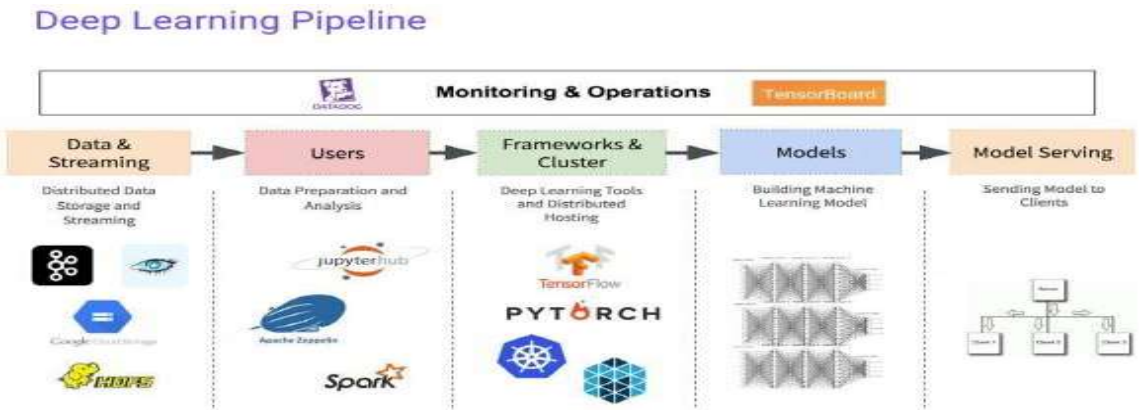


Fig.2 Deep learning pipeline

V. CHOOSING A DEEP LEARNING FRAMEWORK

Using Python and NumPY to implement algorithms to learn deep from scratch is a good way to understand what these deep study algorithms really do. However, it is not increasingly practical, at least not for most people like us. (CNN) [14][15] or revenant neural networks (CNN) or such complicated models because the sensible neural network, implement everything manually from scratch. Today, there square measure several deep learning frameworks that build it straightforward for you to use the neural network, and here square measure a number of the leading.

Fig.3 Deep learning Framework

These Frameworks Of Each Has A Dedicated User And Developer Community And I Think Each Of These Frameworks Is A Credible Choice For Some Subset Of Applications. However, When I See the Below Graph Our Obvious Choice Goes For **TensorFlow.**

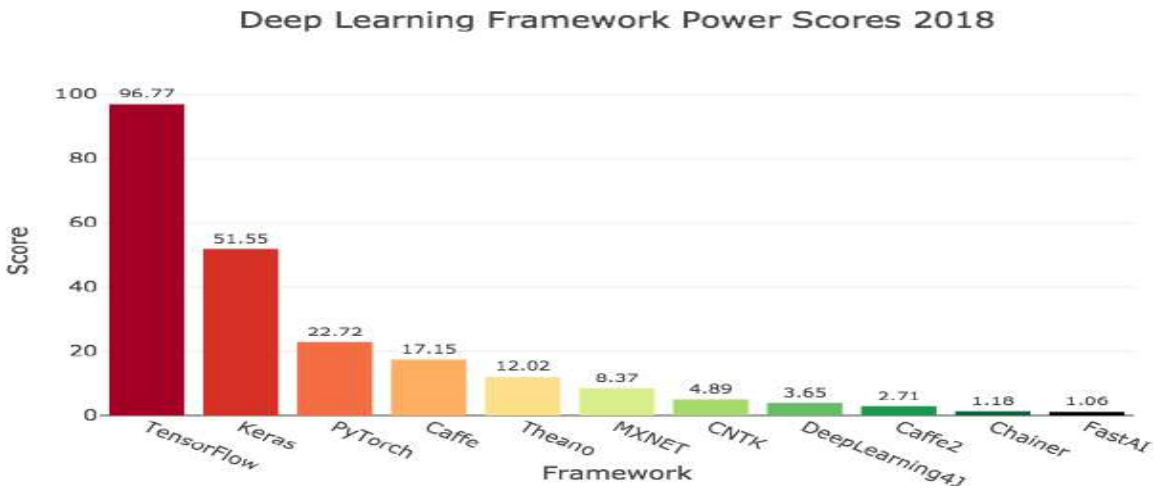


Fig .4 Deep learning framework power score

Rather than strongly endorsing any of these frameworks, we would like to share three factors that Stanford Professor Andrew Ng.[12][13] considers important enough to influence your decision.

1) Ease of programming
This involves developing, repetitive and, ultimately, deploying your neural network into a product where millions of users can use it.

2) Running speed
Training on large data sets can take a long time, and differences in the speed of training between frameworks can make your workflow more time efficient.

3) Openness
Of course, an open structure must of course be open source, but it must also be well managed. Therefore, it is important to use a company framework that we can trust.

VI. DEEP LEARNING FOR MISINFORMATION DETECTION

Misinformation is defined as an observation that deviates greatly from other observations and from there raises the suspicion that it has been generated by a different mechanism. ,[8] It seems that a similar problem has been solved by several techniques. Although many techniques are being used to detect misinformation in social network data, DL is a much better way to use it. However, similar misinformation problems have been solved with different DL techniques. In addition, this type of D.L. Techniques are based on different data characteristics and are used to automatically identify false information. Therefore, we have the following D.L. Technologies are divided into three main categories based on models: (1) discriminatory models, (2) generative models and (3)

hybrid models. All three categories have a large number of architecture models commonly used for MID[17][18]

Techniques for deep learning

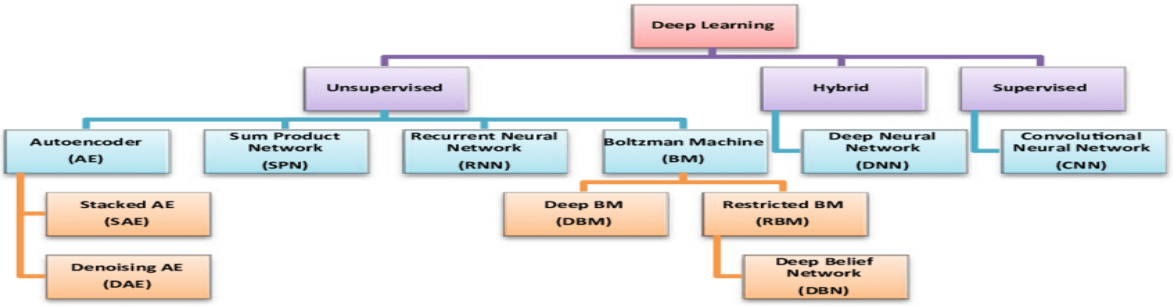


Fig.5 Techniques for deep learning

Recurrent neural networks (RNNs) use sequential information in the RNN network which is essential in many applications where the structure embedded in the data sequence delivers useful knowledge. The advantage of RNNs is the ability to better capture relevant information. To find out rumors, existing methods rely on craft features to use machine learning algorithms, which require huge manual effort. To avoid this issue, RNN to find out the rumors. RNN architecture used for the fake news detection. The authors use different R.N. In the proposed architectures, GRU has achieved the best results in both datasets with an accuracy of0.88 and 0.91, [19][20] respectively. RNN model for learning and diversified the relevant information in posts over time. Additionally, they described that RNNs use sequential information in a network where the structure embedded in the data sequence provides useful knowledge.

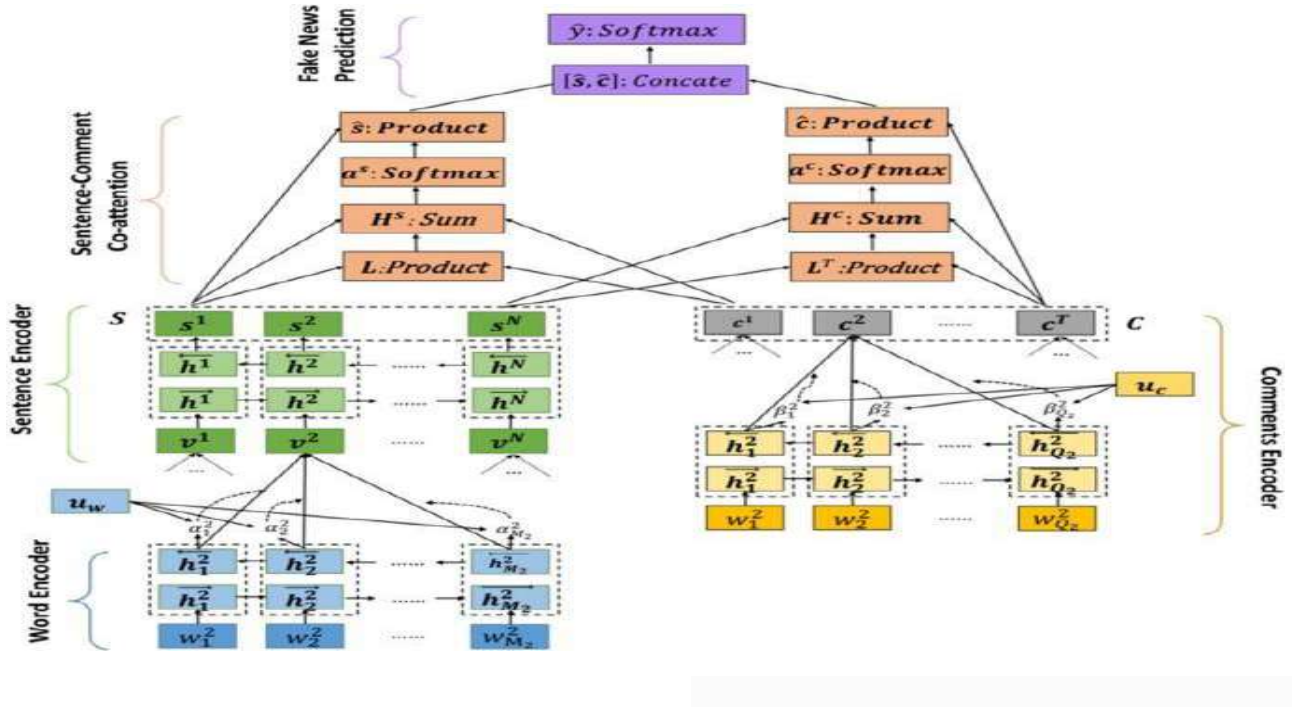


Fig.6 Fake news Prediction

Recursive Neural Network

Researchers are more concerned with identifying unethical users and wants to protect genuine users from fraudulent behavior, therefore, it is the most widely used and successful network for many natural language processing (NLP) functions.. In order to reproduce the pattern of the input layer into the output layer, this network is trained by auto-association. Also, this model analyzes a text word by word and stores the economics of all previous texts in a hidden layer of fixed size, proposing an RVNN architecture to manage the input of different modules. [9] It follows a non- indexed pattern to present a stronger identification of different types of rumored content creations..

In this section, we summarise some limitations which we identified and proposed some ideas to address these limitations:

Semantics Understanding False information that is forged or manipulated to mislead users. It is very difficult for a machine to fully understand such economics. The current study for MID includes a variety of language styles. However, an understanding of the semantic features is required to differentiate between different weapons and improve the performance of MID

.

Multimodal Data for Misinformation In the existing literature, there are many studies related to MID such as detecting rumors, finding fake news, and spam investigations based on multi-modal features. As per the previous study we clarify that misinformation on social media takes the form of text, images or videos and information in various ways can signal for MID.

Content Validation Due to misinformation, users are always confronted with misleading, confusing, controversial issues that need to be addressed very well.. Therefore, in order to easily identify misinformation on social online social media, we need a very good quality fact-checker and a special tool for crowd sourcing content validation can be developed.

Spreader Identification Identifying the influential spreaders in social networks is a very important topic, which is conducive to deeply understand the role of nodes in information diffusion and epidemic spreading among a population

.

False information identification Current research has introduced a variety of misidentification methods. However, most research work focuses on (a) warning users but does not explain why this is misinformation;

(B) Focus more on directly engaged users to find false information. But if users are not upright related, some users play an effective role in spreading misinformation on online social media. As they are not directly related, identifying them is a very difficult job.

Anomalous and Normal User Identification As the number of people relying on online social media increases, dishonest users try to take advantage of this opportunity. Although researchers have used many methods to identify dishonest users, many more approaches can be explored

Bridging Echo Chambers Social media Echo chambers play an important role in spreading the presence of misinformation. One strategy for MID is to bridge conflicting echo chambers so that opposing opponents can be exchanged and taken into account. do research to reduce the effectiveness of polarization.

Mining Disinformation Misinformation can have widespread detrimental social effects. Therefore, mining disinformation is desirable to prevent a large number of people from being affected. However, due to its diversity, complexity, multi-modality and fact- checking costs, it is still not insignificant. Additionally, receiving abundantly labeled data is always unrealistic. The present study argued that due to over-fitting on small labeled datasets, performance is greatly limited. In addition, models learned on one domain may be biased and may not perform well on different target domains. [10]

Future direction

As with anything, there are both good and bad aspects to technological dependence. A lot of research work has been done on MID, and good results have been achieved through various effective techniques. However, we have to keep in mind that the present age is knowledge based and technology based. Therefore, researchers need to think deeply about how their research can transform people's well-being in a technology-based age.[11] Thus, in this paper, we have discussed some of the effective roles that can be used to remove misinformation from DL Moreover, we focused on detecting impact, characteristics, and misinformation using DL techniques. In summary, the following are

- One of the important tasks of DL is that it can work with large-sized data which the other techniques cannot. However, DL also has difficulty to find and process massive datasets, and generally to train the model, DL networks require a lot of time. Most current studies show that researchers can analyze static data on a given topic and predict the positive or negative aspects of that topic. [11]
- The practice of finding false facts from data is very popular and people benefit greatly. However, this includes descriptive research which is not only predictive research, but also explained MID.
- Deep reinforcement learning is a new area of machine learning that enables an agent tolearn in a good interactive environment by experimenting with feedback from it own experience.
- Deep learning faces the over-fitting problem which impacts the execution of the model in real-life situations.

VII. CONCLUSION

Although several researchers argue that the social media and such info obtained from its metrics, could be a key-feature for election prediction, others argue that this approach is simply too oversimplified thanks to the shortage of certainty over the important goal of political discussion on such social medias, as several tend to be satiric and not extremely serious, or the shortage of Associate in Nursing recursive and logic formalism preliminary definitions and even disceptation that the nice performance during this survey, we have a tendency to reviewed varied analysis works on middle in social networks. above all, we have a tendency to took a comprehensive inspect the 5 connected terms of information, rumor, spam, faux news and misinformation and mentioned however information misleads folks on social networks. we have a tendency to additionally mentioned the importance of previous works as these is also useful to alternative researchers World Health Organization would like to explore the sector. Compared to existing search approaches, we've found that metric capacity unit is Associate in Nursing economical and effective technique for measure information on social on-line social networks. we have a tendency to emphasize that DIL is currently the leading technology for finding middle issues because it helps to totally determine false facts. The result and performance ar glorious and it's almost like human performance.

REFERENCES

[1] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, et al. (2015) Tensorflow: large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.

[2] Aiello LM, Petkos G, Martin C, Corney D, Papadopoulos S, Skraba R, Göker A, Kompatsiaris I, Jaimes A (2013) Sensing trending topics in twitter. IEEE Trans Multimedia 15:1268–1282

[3] P. T. Metaxas, E. Mustafaraj, and D. Gayo-Avello, ‑How (Not) to Predict Elections,‖ pp. 165–171, IEEE, Oct. 2011.

[4] R. Ushigome, T. Matsuda, M. Sonoda, and J. Chao ‑Examination of Classifying Hoaxes over SNS Using Bayesian Network,‖ pp. 606–608, IEEE, Nov. 2017.

[5] E. C. Tandoc Jr, Z. W. Lim, and R. Ling, ‑Defining fake news a typology of scholarly definitions,‖ *Digital Journalism*, pp. 1–17, 2017.

[6] S. Kumar and N. Shah, ‑False information on web and social media: A survey,‖ *CoRR*, vol. abs/1804.08559, 2018.

[7] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, ‐Detection and Resolution of Rumours in Social Media: A Survey,‖ *ACM Computing Surveys*, vol. 51, pp. 1– 36, Feb. 2018.

[8] S. Nieminen and L. Rapeli, ‐Fighting misperceptions and doubting journalists objectivity: A review of fact-checking literature, *Political Studies Review*, p. 1478929918786852, 2018.

[9] L Zheng and C. W. Tan,‐A probabilistic characterization of the rumor graph boundary in rumor source detection,‖ pp. 765–769, IEEE, July2015.

[10] J. A. Ceron-Guzman and E. Leon-Guzman, ‐A Sentiment Analysis System of Spanish Tweets and Its Application in Colombia 2014 Presidential Election,‖ pp. 250–257, IEEE, Oct. 2016.

[11] J. Radianti, S. R. Hiltz, and L. Labaka, ‐An Overview of Public Concerns During the Recovery Period after a Major Earthquake: Nepal Twitter Analysis,‖ pp. 136– 145, IEEE, Jan. 2016.

[12] S. Ahmed, R. Monzur, and R. alit, ‐Development of a Rumor and Spam Reporting and Removal Tool for Social Media,‖ pp. 157–163, IEEE, Dec. 2016.

[13] Alkhodair SA, Ding SH, Fung BC, Liu J (2020) Detecting breaking news rumors of emerging topics in social media. Inf Process Manag 57:102018.

[14] I. Y. R. Pratiwi, R. A. Asmara, and F. Rahutomo, ‐Study of hoax news detection using nave bayes classifier in Indonesian language,‖ pp. 73 78, IEEE, Oct.2017.

[15] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, ‐Fake news detection on social media: A data mining perspective,‖ *SIGKDD Explorations*, vol. 19, no. 1, pp. 22–36, 2017.

[16] V. Subrahmanian, A. Azaria, S. Durst, V. Kagan,A. Galstyan, K. Lerman, L. Zhu, E. Ferrara, A. Flammini, and F. Menczer, ‐The DARPA Twitter Bot Challenge,‖ *Computer*, vol. 49, pp. 38–46, June 2016.

[17] C. Shao, G. L. Ciampaglia, O. Varol, A. Flammini, and F. Menczer, ‐The spread of fake news by social bots,‖*arXiv preprint arXiv:1707.07592*, 2017.

[18] S. Vosoughi, D. Roy, and S. Aral, ‐The spread of true and false news online,‖ *Science*, vol. 359, no. 6380, pp. 1146– 1151, 2018.

[19] Q. Wang, T. Lu, X. Ding, and N. Gu, ‐Think twice before reposting it: Promoting accountable behavior on Sina Weibo,‖ pp. 463–468, IEEE, May 2014.

[20] K. Wu, S. Yang, and K. Q. Zhu, ‐False rumors detection on Sina Weibo by propagation structures,‖ pp. 651–662, IEEE, Apr.2015

Sign Language Communication using Convolution Neural Network

Gunjan Chugh[1], Ankit Sharma[2], Saijal Gupta[3], Monica Batra[4]

[1,3,4]*Assistant Professor, Department of Information Technology, ADGITM, Delhi*
[1]chugh.gunjan8917@gmail.com
[3]saijalgupta92@gmail.com
[4]monica.batra88@gmail.com

[2]*Systems Engineer, Infosys Ltd, Jaipur*
[2]ankitbhardwaj4903@gmail.com

***Abstract*- In this paper we present a Sign Language(SL) Communicator for speech and hearing-impaired peoples. Communication is an important aspect of the society, when it comes to speech and hearing-impaired people everyone just relies on sign language. The proposed communicator is able to provide a better way of communication between speech/hearing impaired people and the people who can speak and hear but don't understand sign language. The SL communicator has two phases, the first phase is to convert sign language into text and the converted text is then translated into audio for hearing people, and in the next phase we convert speech to sign language for speech and hearing-impaired people. The conversion of sign language into text is done by using CNN (Convolution Neural Network) and the conversion of text to speech and vice-versa is done by using speech recognition. The main advantage of this communicator is that it is built for a system having average specification so that it will function on all most all the systems.**

***Keywords*- Convolution Neural Network, Deep Learning, Sign Language Recognition**

I. INTRODUCTION

Communication is an integral aspect of interaction between peoples and when it comes to deaf community, they rely only on written and sign language communication. Sign language consists of different gestures and hand motions combined with facial expressions or body movements [1]. The Greek philosopher Socrates, who lived in 469-359 B.C., is named in honour of sign language. He analysed how people could interact if they didn't have a voice or a tongue. [2]. Dumb deaf people use hand motions or gestures to convey their feelings [1]. Although written communication can be possible between hearing and deaf peoples but it is very difficult, because the Deaf community is known to be less skilled in writing a spoken language. Also written communication not practical at all consider the following example, when a person (who is deaf) got lost then you need to communicate quickly with people around you and it is not necessary that the person you encounter understands sign language. A sign language could be used to express characters, words, or sentences by using various hand gestures. Hand gestures are particularly useful for communicating any word or feeling. As a result, despite the development of writing conventions, people all over the world use hand gestures to express themselves. [3].

The main objective of this work is to provide a better and more practical way of communication between Deaf and Hearing community. There are two phases in building the live communicator. The first phase is to train a Convolution Neural Network on hand trained images of 96 × 96 pixels, the images are grayscale images. The second phase is to perform text to Speech Conversion and vice-versa using Google Text To Speech API.

II. RELATED WORK

Sign Language Recognition deals with higher number of classes and this involve very similar kind of features, but the basic approach is still related to gestures recognition. Mitra, S., & Acharya, T. [4] proposed a gesture recognition system involving hand gestures and facial expression. They took a different approach instead of using CNNs they use particle filtering optical flow and hidden Markov models. Kang, B. [5] proposed a Sign Language recognition system using depth map. The system is built using CaffeNet which is a very similar implementation of Alex Net, they use this model to classify 31 classes (Signs).

When we talk about Sign Language Recognition the signs vary with languages, L. Pigou et al. [6] proposed an Italian Sign Language recognition system which uses the Microsoft Kinect, convolution neural networks (CNNs) and GPU acceleration. The proposed system is able to classify 20 different Italian gestures with a cross – validation accuracy of 91.7%.

Chinese Language is considered very difficult to pronounce as it contains large amount of words so it's respective interpretation in sign language is also difficult. Chai. et al. [7] proposed a system similar to L. Pigou's, here they used Microsoft Kinect to recognize 239 words of Chinese Sign Language (CSL). Besides, use of static gestures recognition they also constructed a system

for tracing the trajectory of 3D movement of hands. Afterwards this path trajectory is then aligned with a gallery of known trajectories.

More enhanced network was done in [8] for building a vision based human recognition system. He used a more complex convolution neural network which deals with dynamic gesture recognition.

III. METHOD

A. Dataset

We obtained the dataset by capturing images of our own hand trained gestures (This gestures includes 27 classes which contain 26 English alphabets and 1 for empty frame). The dataset contains about 1300 ~ 1400 images per class, in total around 36636 (approx.). This images shown in Figure 1. are flipped 180 degree vertically as we have used the inbuilt webcam of our system so the images which are captured were the exact mirrored images of the actual hand gestures. We choose a nominal resolution for the images which is 96x96, the reason behind choosing this resolution is that it is small enough to process all the images and large enough to extract the feature map.

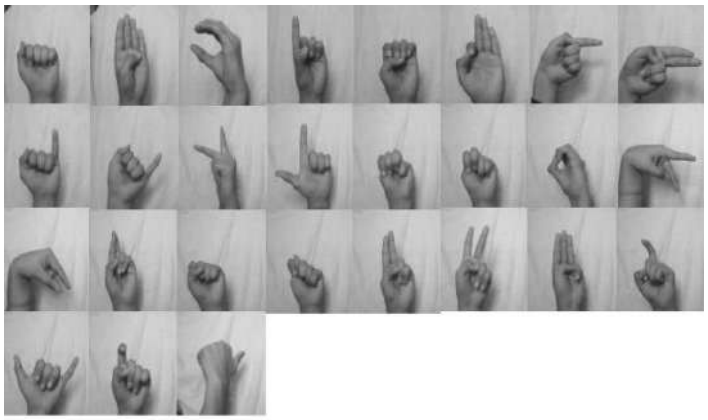


Fig. 1 Some examples from pre-processed dataset ranging from A-Z and last symbol for space.

B. Data Pre-processing

For Data pre-processing, we converted the coloured images into grey-scaled images this is done by launching web camera and accessing the output frame by frame, this individual frame is first converted into grayscaled (reducing the dimension of the image from 96x96x3 to 96x96 shown in Figure 2(a) and 2(b)) after this the frame is flipped 180 degree vertically (shown in Figure 2(c)) and were simultaneously saved in the specified directory [9].

The next step in Data pre-processing is to manually remove some initial frames in which the gesture is not perfectly captured within the region of interest, this manual reduction is necessary as it will improve the overall accuracy of the model.

(a)Original cropped image

(b) Converted grey-scaled image (c) Flipped image

Fig. 2 Steps involved in Data pre-processing.

C. Convolution neural network

CNNs [9] are very powerful tool when it comes to process or classify complex data like image voice etc. The conventional ANN (Artificial Neural Network) can also be implied on this sign language recognition system but then the input nodes will be much more than the CNN (9216 Node as compared to 1024 nodes in CNN) this is the major reason why the industry leaders like Google, Facebook and Amazon use CNNs for image recognition. And researchers at Google applied CNNs on video data[10].

Now let's discuss working of CNN analogous to human brain, CNNs are similar to the visual cortex of the human brain. The artificial neuron in the network will attached to a local group of the visual field, called a receptive field. This is achieved in artificial CNN by performing discrete convolutions on the image with filter values as trainable weights. Considering the type and amount of images, multiple filters are used in different layers and together with activation functions of the neurons, they form feature maps. Further Pooling layers are used to extract only useful features and discarding the rest resulting in a smaller sized network as shown in Figure 3.

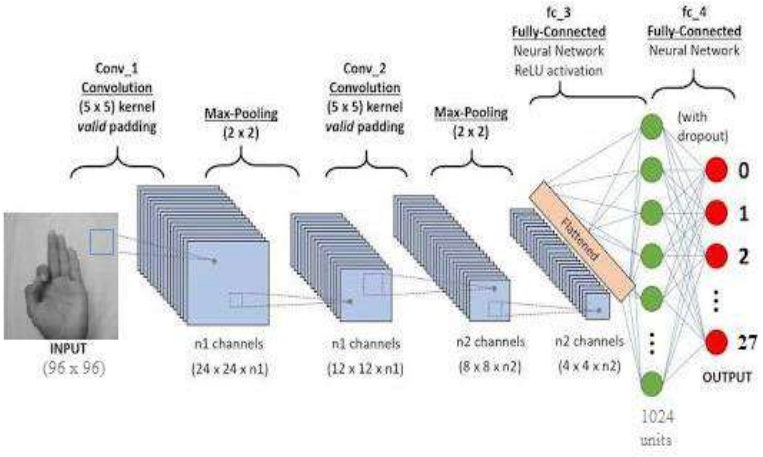


Fig. 3 Architecture of the Neural Network [6].

It is not wise to have large number of nodes (neurons) in a layers as compared to the numbers of layers. A network with one layer having large number of nodes may overfit the data and also it increases the complexity that's why we use a deep layered network instead of network having more neurons in a single layer.

D. Required architecture

As discussed earlier, the pooling layers used in this layer are max-pooling (In max-pooling the maximum value in the region remains in the feature map). The max-pooling layers used in the network are of size 2x2 as shown in Fig. 3. Although 3-D pooling layers can be used in the network (3-D max-pooling layers [6]) but we achieve better results by using 2-D pooling layers.

The architecture consists of eight layers, first layer is the input layer, then the next five layers are the convolution layers with respective max-pooling layers [11], the second last layer is the classical ANN layer, which is only one layer deeper. This layer

provides the classification and the last layer is the output layer which provides the results for 27 different classes. All the artificial neurons used in the first seven layers have rectified linear activation function (ReLUs, [12]), and the last layer uses Softmax activation function. Further we have to convert the predicted output to its respective audio profile and vice-versa. By using Google's built in API for speech recognition we convert the output of the network to speech, since the speech conversion is done by Google's API the system should be connected to the internet all the time.

In the second phase we convert the speech input from hearing people into the respective signs, this is done by the same process as discussed earlier but the speech to text involves one additional step. The output of speech to text conversion is then matched with the signs which were stored in a dataset, this dataset consists of images with three tags.

E. Optimization and training

To reduce overfitting in the network dropout is used with a constant value of 0.8 [8]. For better results in less time we used Adam optimizer, though SGD (Stochastic Gradient Descent) Nesterov optimizer can be used to achieve higher accuracy, but we train this model only on CPU that's why we use Adam. Random initialization is used for the initialization of the weights.

This model is trained on our single machine with Intel core i3-5005U dual core processor clocked at 2.0 GHz each, and 4GB of DDR3 RAM. The whole model is built on Python with tflearn and opencv (used for data capturing), speech-recognition and gTTS is used for sign to speech conversion and vice- versa.

IV. RESULT

In the end we got a pretty good model with test set accuracy of 92.67% and the training set accuracy of 96.7%, this is achieved USING a low powered, low end system without having any external GPU, and the images used for training are single channel grey-scaled images.

V. CONCLUSION

By this work, we represent a useful implementation of CNN to build a model which recognises the sign language with 27 different classes having an accuracy of 92.67%. The mentioned model can recognise different signs very accurately. Further this can be improved by incorporating a higher resolution image dataset and a good hardware.

REFERENCES

[1] D. Dhake, M.P Kamble, S.S. Kumbhar & S. M Patil, "Sign Language Communication with Dumb and Deaf People" *International Journal of Engineering Applied Sciences and Technology*, Vol. 5, Issue 4, ISSN No. 2455-2143, Pages 254-258, 2020

[2] B.P Bala, & L.A Song, "Android app for improvising sign language communication in English and Hausa", *International Journal of Advances in Scientific Research and Engineering*, 6(2), 15-24, 2020

[3] A. Wadhawan, & P. Kumar, "Deep learning-based sign language recognition system for static signs", *Neural Computing and Applications*, 32(12), 7957-7968,2020

[4] S. Mitra, & T. Acharya, "Gesture recognition: A survey. IEEE Transactions on Systems, Man, and Cybernetics", *Part C (Applications and Reviews)*, 37(3), 311-324,2007

[5] B. Kang, S. Tripathi, & T. Q Nguyen, "Real-time sign language fingerspelling recognition using convolutional neural networks from depth map", In 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), pp. 136-140. IEEE, 2015.

[6] L. Pigou, S. Dielman, P.J Kindermans, B. Schrauwen, "Sign Language Recognition using Convolution Neural Networks, European Conference on Computer Vision" 572-578, 2014

[7] X. Chai, G. Li, Y. Lin, Z. Xu, Y. Tang, X. Chen, & M. Zhou, "Sign language recognition and translation with Kinect" In IEEE Conf. on AFGR ,Vol. 655, p. 4, 2013.

[8] R. Poppe, "A survey on vision-based human action recognition. Image and vision computing", 28(6), 976-990, 2010

[9] S. Escalera, X. Baró, J. Gonzalez, M.A Bautista, M. Madadi, M. Reyes & L. Guyon, "Chalearn looking at people challenge 2014: Dataset and results" In European Conference on Computer Vision (pp. 459-473). Springer, Cham., 2014

[10] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar & L. Fei-Fei, "Large-scale video classification with convolutional neural networks. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition", pp. 1725-1732, 2014.

[11] J. Isaacs & S. Foo "Hand pose estimation for american sign language recognition. In Thirty-Sixth Southeastern Symposium on System Theory", Proceedings of the (pp. 132-136). IEEE, 2004

[12] X. Glorot, A. Bordes, & Y. Bengio, "Deep sparse rectifier neural networks. In Proceedings of the fourteenth international conference on artificial intelligence and statistics" (pp. 315-323). JMLR Workshop and Conference Proceedings, 2011

Algorithm Visualizer (ALGOVIZ)

[1]Archana Walecha, [2]PiyushJain, [3]RohanSinghBisht, [4]ShrutiRustagi,[5]Lalita Luthra
[1,2,3,4,5]Department of Computer Science Engineering
[1,2,3,4,5]*Dr. Akhilesh Das Gupta Institute of Technology and Management, Shastri Park, Delhi*

[1]archana999walecha@gmail.com
[2]piyushjain1555@gmail.com
[3]rsb97695@gmail.com
[4]rustagishruti016@gmail.com
[5]Lalita.luthra@adgitmdelhi.ac.in

***Abstract*- Acquiring knowledge about algorithms and programming skills is a difficult and complex process in particular. Every time we had to visualize the pattern in which the algorithm is performing and for complex data structures involving many transitions in a single step it becomes more hectic and confusing at the same time. Hence our primary goal is to facilitate student's learning, understanding of the concepts and skills development. Algorithm visualization and can be defined as the use of continuous images to convey some useful information about algorithms. That information can be a visual illustration of an algorithm's operation, of its performance on different kinds of inputs, or of its execution speed versus that of other algorithms for the same problem. Within the paper we discuss the possibility of enriching the standard methods of teaching algorithms, with the algorithm visualizations. As a step in this direction, we introduce the AlgoViz, algorithm visualization platform based on reactjs to see the algorithms in action.**

I. INTRODUCTION

Algorithms and data structures as an essential part of knowledge in a framework of computer science have their stable position in computer science curricula, since every computer scientist and every professional programmer should have the basic knowledge from the area [1]. With the increasing number of students in the current higher education systems in last decades, introduction of appropriate methods into the process of their education is also required. Our scope here is the higher education in the field of computer science. So, within the paper, we discuss the extension of standard methods of teaching algorithms that is using the whiteboard or slides, with the algorithm visualizations. According to [2] they can be used to attract student's attention during the lecture, explain concepts in visual terms, encourage a practical learning process, and facilitate better communication between students and instructors. Interactive algorithm visualizations allow students to experiment and explore the ideas with respect to their individual needs and pace.

Except the algorithm visualization, the term software visualization is also often used within the papers published in last years. It usually covers both visualization of algorithms and visualization of data structures, but sometimes also another aspects of software (like its development process) are considered, too [3]. Algorithm visualization, as part of software visualization, could be described as "graphical representation of an algorithm or program that dynamically changes as the algorithm runs"[4].

Modern approaches to software visualization were brought in the2000's by the introduction of system BALSA (Brown & Sedgewick, Brown University, USA) [5]. Some of contemporary solutions include systems like TRAKLA , ANIMAL [6], JAWAA or Algorithms In Action. Concise overview of development in the area of software visualization we provided in [7], so it is not our intention to analyse this topic within the paper.

In this project, we have made a platform for sorting algorithm visualization. It currently includes six major sorting algorithms namely bubble, insertion, selection, merge, quick and heap sort. The project is divided into mainly two parts: one is the react visualizer that uses several animation and timeout functions , second one is the homepage that gives the idea about the working of the algorithm, their complexities and code.

II. ALGOVIZ (Algorithm Visualization Platform)

Based on analysis of existing solutions, we decided to start developing our own algorithm visualization platform named AlgoViz. The motivation behind the decision is that the platform is intended to be used as a support tool within the subject

Data structures and algorithms.

As a development platform for the project was selected Java, ensuring high portability and very good support by available tools, libraries, etc. Another important decision to made was the selection of software framework to support extensibility.

A. Architecture of the Platform

Basically, we can think of the AlgoViz application as consisting of two cooperating parts: the main page i.e. home page and a set of independent algo modules. The algo module consists of several functions providing support for controlling the algorithm execution and rendering the algorithm visualization. The home page is built with pure HTML and CSS and is further integrated with Reactvisualizer.

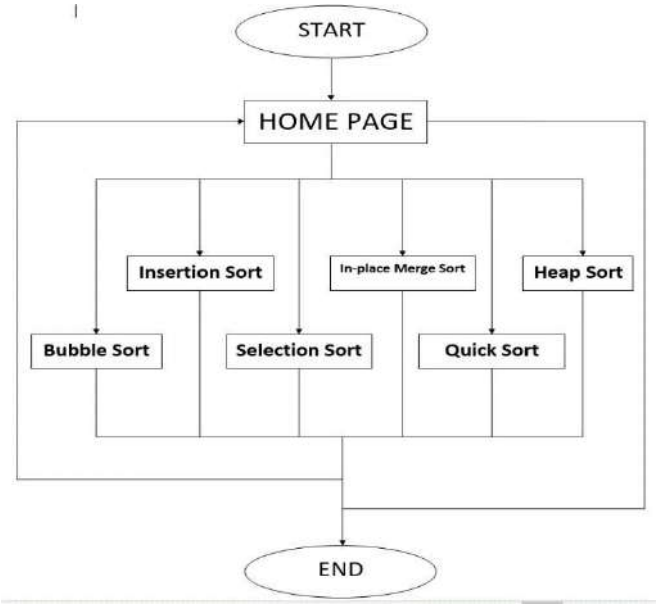


Fig 1. The structure of the home page

The algo module, on the other hand, contains the code of algorithm to be visualized, the animate functions and main class to render the entire webpage written in HTML5. It can be further divided into the two modules if needed into the visualization part(reactjs) and the layout whose role is to support and display the content related to the algorithm. Given below is the flowchart for the visualization page using render() function and animateSort() which is user defined.

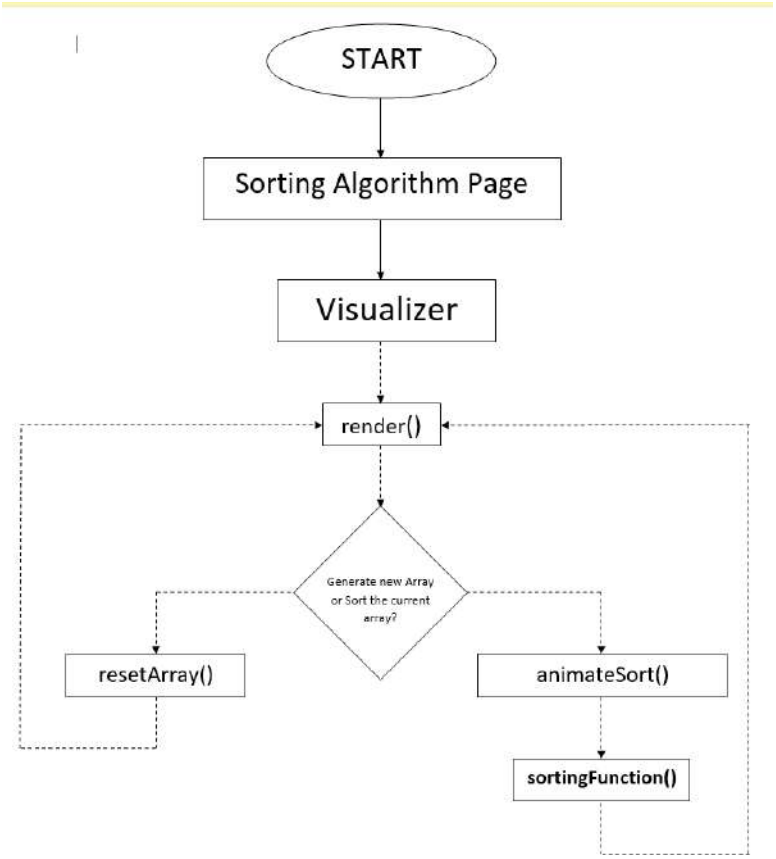


Fig 2. Structure of the algo module

B. Developing the algo module for thePlatform:

Algo module development starts by creating the interface with the name of algorithm to be visualized (e.g. MergeSort). The interface just created will contain

1)*Sorting algorithm function:* The mergesort function have one additional parameter animations which is passed to carry out the bar animations. The algorithm will be called on given array and return animations statement will be used instead of a return statement for visualization purposes. Every time there is a swap statement the elements are pushed in animation using animations.push([true,i,j,array[j]])statement.

```
functionmergeSort(array,start,end,animations)
{
      if(start !==end)
      {
            varmid=Math.floor((start+end)/2);
      mergeSort(array,start,mid,animations);
      mergeSort(array,mid+1,end,animations);
      return(merge(array,start,mid,end,animations));
      }
      else{
```

2)*Function to animate the bar graph representing the array: In* this function, the mergesort() function is called to get animations. Then, the bar values and styles are defined and setTimeout() function is used to make the delay and hence setting speed dynamically.

```
exportfunctionmainSort(array,speed)
    {
        var animations = [];
        animations=mergeSort(array,0,array.length-1,animations);
    constarrayBars=document.getElementsByClassName('array-bar');
    consttext=document.getElementsByClassName('bar-value');
    while(i<animations.length)
        {. . . . . .. .... . . . .}
    }
```

3)*Drawer function for displaying code :*Additional function required for tab switching between the available languages(C++, Java, Python, Javascript) for the sorting code.(Partof rendering the page)

```
 functionopenTab(event,codeName){ vari,
tabcontent, tablinks;
   //Getallelementswithclass="tabcontent"andhidethem
tabcontent=document.getElementsByClassName("tabcontent");
for(i=0;i<tabcontent.length;i++){
        tabcontent[i].style.display="none";
    }
}
```

4)*Main class to render the entire page: The* class merge extends the component class from react and is responsible for rendering the page. setArraySize() function creates the array of the size given by the user. setSpeed() function is used to slow down the speed of animation. resetArray() will generate a new random array of the same size. Atlast, the render() function will display the landing page and onClick it will call the merge()function.

```
    classMergeextendsComponent{
    componentDidMount(){
        . . . . . }
        handleChange(event){
            this.setState({size:event.target.value});
        }
          handleSpeedChange(event){
      this.setState({speed:event.target.value});
        }
        setArraySize(){
    } setSpeed(){   }
    resetArray(){       }
        render() {
```

```
                /*HTML5 layout*/
            }
}
```

C. Output:

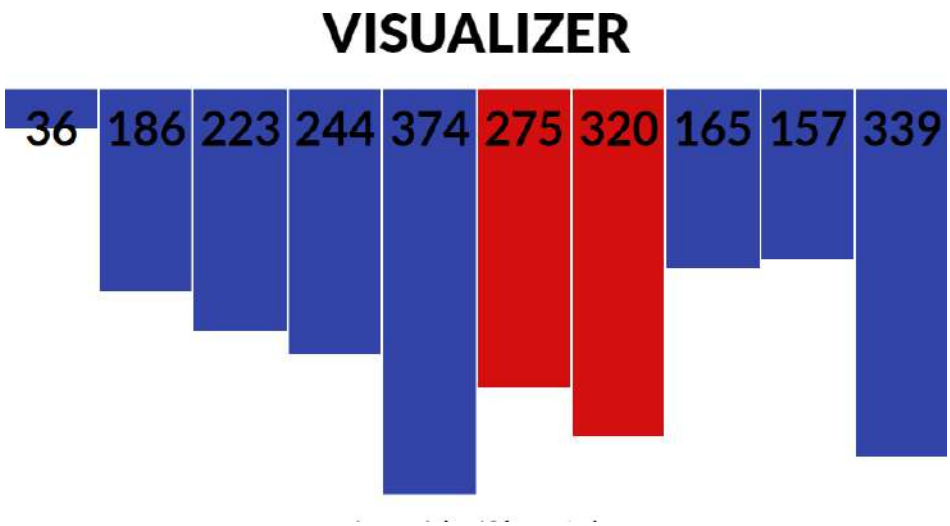


Fig 3. Array is being sorted

In this the user can enter either their own values or random array will be generated as shown above. The array will be sorted using the merge sort algorithm where the array is divided into parts and then merge in a sorted array. The red values are being compared. After that the array will be sorted.

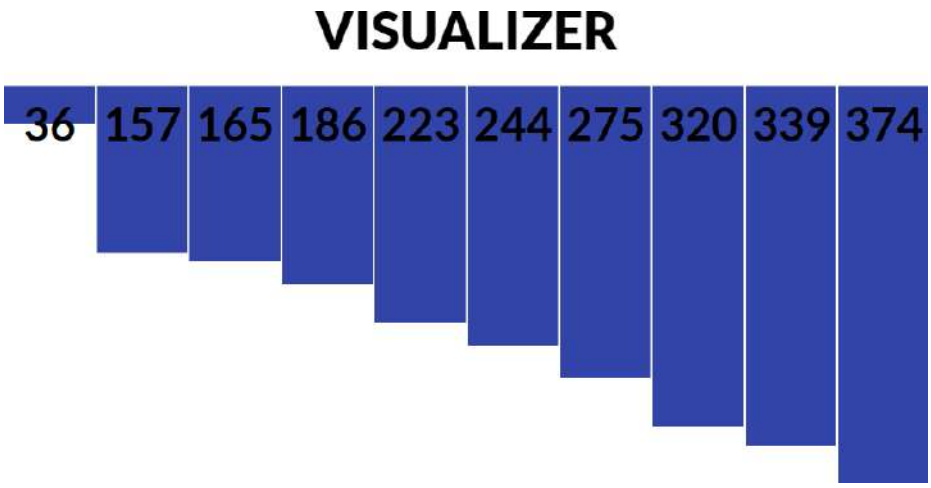


Fig 4. Sorted Array

1) *Buttons used and their functioning:*

1. 

Generate random array button will create the random array of size entered by the user. On click, this button will trigger the resetarray() function. This function will create array bars of array values which will be

created using random function. If the array size is maximum, then number of array will be shown according to the window size.

2. 

Enter Array fill create an input box to allow user to enter array.

3. 

Confirm Array will allow user to enter array of whichever array size they want. On click, this button will trigger the createarray() function. This function will get the values from user and then split it using split function and then convert it into int using parseInt function. It will create the bars according to the array given by the user.

4. 

Bubble Sort will sort the array values. On click, this button will trigger animateMergeSort() function. This function consist two parameters, array values and speed. Speed parameter is used to see how the values are swapping and value of speed is entered by the user. This function is different for different algorithm.

III. TESTING

Initially all the four algorithms were implemented together only by selecting the type. It was tested for the randomly generated array with different time delays and the results were accurate. Then all the sorting algorithms are divided into their own modules. Then the testing was done for different sizes of array depending upon the width/screen size of the device. The outcome is : minimum size of the array was 8 and the maximum limit is varying. In case of a laptop the maximum size is 192(varying) and for phone it is 53(varying). Then both kind of modules : visualizer(reactjs) and the webpages(HTML)are integrated. The results were okay as values are in rem and it was working properly. At last, we included the input array option to take input from user and working upon that.

IV. FUTURE SCOPE

Currently this includes only the sorting algorithms and only sorting algorithms are visualized. In future we will try to add other sorting algorithms also like path finding algorithms and so on. We will try to add stop and pause buttons so that user can understand how the values are swapping. We will try to add that which part of the code is running for better understanding.

REFERENCES

[1] K. Mehlhorn, P. Sanders, Algorithms and Data Structures (Springer-Verlag, Berlin Heidelberg, 2008)

[2] S. Khuri, Designing Effective Algorithm Visualizations, In proceedings of: First International Program Visualization Workshop, ITiCSE.

[3] S. Diehl, Software visualization: Visualizing the Structure, Behaviour, and Evolution of Software (Springer, NewYork, 2007)187

[4] M.E. Tudoreanu, R. Wu, A. Hamilton-Taylor, E. Kraemer, Empirical Evidence that Algorithm Animation Promotes Understanding of Distributed Algorithms, In proceedings of: IEEE Symposium on Human Centric Computing Languages and Environments, HCC02,Arlington, Virginia

[5] M.H. Brown, R. Sedgewick, A system for algorithm animation, Proceedings of the 11th annual conference on Computer graphics and interactive techniques,SIGGRAPH'84

[6] G. Rößling, B. Freisleben, ANIMAL: A system for supporting multiple roles in algorithm animation, J. Visual Lang. Comput. 13(3),341–354

[7] S.Šimonák,AlgorithmVisualizationUsingthe VizAlgo Platform, ActaElectrotechnica et Informatica 13(2),54–64,2013

Health Diagnostic System

Gitansh Gera[1], Himanshi Jindgar[2], Saanvi Jain[3], Sachin Parashar[4], Mr. Amit Kumar Pandey[5]
[1,2,3,4]*Students,* [5]*Assistant Professor*
Department of Information Technology
[1, 2,3,4,5] *Dr. Akhilesh Das Gupta Institute of Technology & Management,*
New Delhi, Delhi-110053, India
[1] gitanshgera10@gmail.com
[2] honeyjindgar@gmail.com
[3] saanvijain1999@gmail.com,
[4] sachinparashar52@gmail.com
[5] amit.pandey@adgitmdelhi.ac.in

***Abstract*— A Health Diagnosis System is an act of identifying whether a person is suffering from a disease by its symptoms. The best way of treating a disease is to determine it at an early stage. Patients suspected of any kind of chronic disease are mostly recommended to be diagnosed by expensive technologies such as MRI. This method is very resource-intensive. In an attempt to overcome this resource-intensive procedure, this project has created programs that simulate expert human reasoning. In this research paper, we have studied how data mining plays an efficient role and helps in accurately predicting medical data and helps in early patient care.**

***Keywords*— Machine Learning, Chronic Diseases, Data Mining, Disease Diagnosis, Deep Learning**

I. INTRODUCTION

Any harmful deviation of a person from standard behaviour either mentally or physically, which can be associated with certain symptoms, is known as a disease. Fundamentally disease represents a condition of breakdown from the psychological balance of the body. Diagnosis is required to check whether or not a person is suffering from a certain disease. A medical diagnosis examines the state of the person's symptoms and concludes which disease the person is suffering from and distinguishes it from other possible conditions. Disease diagnosis carries out a very principal role in the healthcare industry and the study of the normal state of human beings becomes extremely vital to diagnose a disease or check its symptoms. The information required from medical diagnostics is typically collected from an individual with the bits of help of various medical tests.

Diagnosis is often a challenging activity because it requires complex reasoning and information to determine the health status of a person. Various reports in medical science suggest that every person faces one health diagnosis error in his lifetime and it can harm the health of the patient by not providing him with appropriate treatment. The patient holds the key information and is central to the solution required in diagnosis, any ambiguity in the data can lead to devastating consequences. In this paper, we study machine learning algorithms that assist us in quick diagnosis by developing a model which very efficiently and fastly analyzes data and delivers results faster. Preparing the model using unstructured data would help us in developing the accuracy of the model and prediction would be done using KNN, Decision Tree.

Machine learning helps us to optimize the performance of a model by using a predefined data set. Machine Learning has a great impact on the healthcare field, science, and technology. Data mining as a part of machine learning assists us to draw a pattern or a relationship between databases. It is relatively new and promising technology helps to forecast the future styles or patterns in an event and summarize it into relevant information which can be used to find new patterns and trends between the variables by digging deep into the medical data. At the healthcare level, it aids in predictive machine learning which leads to effective treatment of the patient and helps in managing health care at various levels.

II. RELATED WORKS

Through this section, we will discover the existing works that have been already done in this field. The early works in data mining are credited to K Vembdaswamy et al.[1], it suggested that the majority of data that the health industry collects are not mined and a huge amount of unmined data gets collected in pursuit of finding hidden patterns and information. Awotunde et al. [2], in their research medical diagnosis using fuzzy logic, recommended the development of a medical diagnosis system using fuzzy logic to increase the accuracy of the project; it also aids in drug prescription and record-keeping of the patient. Simarjeet Kaur et al. [3] analyzed the AI techniques which were previously used in the Medical Diagnostic System.

Regression as a technique in health diagnosis was studied by Van Mourik et al. [4]. It surveyed how automatic diagnostic techniques based on machine learning algorithms are more efficient and helpful than manual surveillance. Bhattacharjee et al. [5] and Sinha et al. [6] studied an illness that is caused by a bodily response to an infection commonly known as sepsis.It occurs when chemicals are released within the bloodstream to fight an infection that triggers inflammation throughout the body. The authors studied various sepsis screening techniques and highlighted which one is better used in present hospital conditions. Another study in sepsis reported obstacles in blood collection methods in which they discussed how modern machine learning techniques can take care of these minor drawbacks in sepsis detection and treatment.

III. RESEARCH METHODOLOGY

A.DATA COLLECTION

Collecting data is the first significant step towards the real development of a machine learning model. For data collection previous works, texts and research would be referred alongside using secondary data (direct consultation from general practitioner) collection method and data mining. Data collection constitutes a fundamentally critical step as the more and better the data we will collect, the better will be working and the performance of our model. In this model, we diagnose 3 diseases and have three different data sets for each one. The heart disease dataset consists of 303 records and 13 attributes, the breast cancer data set contains 569 entries and 32 columns, whereas the chronic kidney disease data set contains 400 entries and 26 columns. There is no missing value in the data set and we display an attribute called target in the dataset which tells whether the person suffers from the disease or not. It takes various attributes into account such as maximum heart rate, exercise-induced angina, no of major vessels, ST-segment. The structural data contains the symptoms of the patient while the unstructured data comprises textual format.

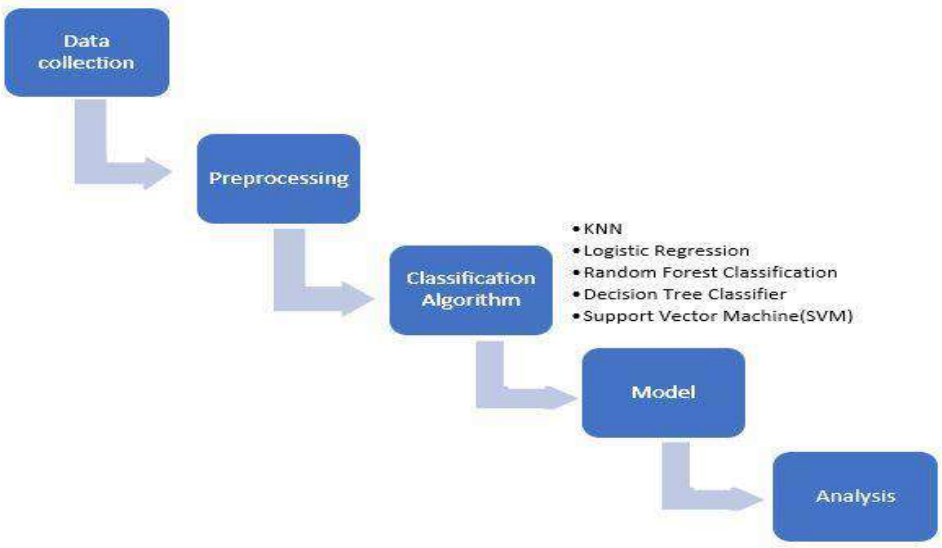


Fig. 1 TIMELINE OF MODEL

B.PROPOSED MODEL

A Web application is developed which will predict whether a person has heart disease, chronic kidney disease, breast cancer or not, based on medical inputs provided by the users. It accepts the structured and textual sort of data as input to the machine

learning model. It will predict the disease based on symptoms and various test results and the model is deployed to the web using Flask, React JS, and Node JS.

IV. MACHINE LEARNING ALGORITHM

The disease prediction system is implemented using the five data mining algorithms i.e. Decision tree classifier, random forest classifier, support vector machine(SVM) classifier, KNN, and logistic regression. The description and dealing of the algorithms are given below.

A. *KNN (K- Nearest Neighbors)*

K-NN Is built on a supervised learning technique and is also called a lazy learn algorithm because it stores the dataset and performs the action on the dataset at the time of classification. It is used both for regression and classification. The K-NN algorithm is case-based. It stores all data points based on similarity and categorizes the new data rested on its availability[7]. A case is being classed by a majority vote of its neighbors, and the case is assigned to the class which is most common amongst its K nearest neighbors measured by a distance function. It is one of the easiest and versatile algorithms to be used in the healthcare system. It is used to propose a system to classify which disease will happen based on symptoms.

B. LOGISTIC REGRESSION

Regression can be defined as a machine learning method that allows us to predict a continuous variable that varies by multiple predictor variables.Logistic Regression is employed when the variable is categorical. It is a type of predictive analysis that uses maximum estimation as a method of approximation. It was exercised in biological science in the early twentieth century after that it was used in many social sciences applications.

C. Decision Tree

A decision tree may be a graphical representation for obtaining all the possible solutions to a drag supported by the given set of conditions. It starts from the root node and further expands into a tree-like structure, decision nodes are used to carry out decisions while leaf nodes are used for outputs. It is a supervised learning technique mostly used for classification problems. It accepts a straightforward choice derived from information highlights. The features of the given dataset are stored in internal nodes, and tests are performed based on these features [8].

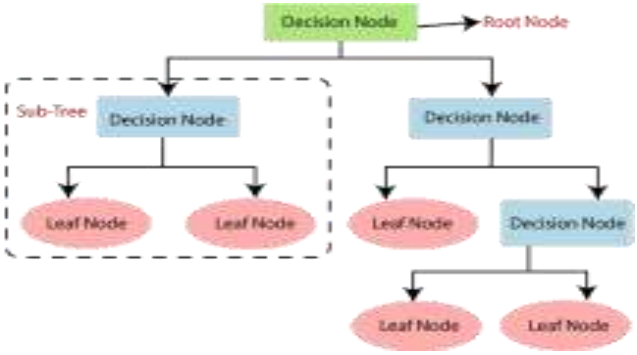


Fig. 2 REPRESENTATION OF A DECISION TREE

D. Random Forest

Random forest contains a substantial number of decision trees on various subsets of the given dataset. It is often used both for classification and regression problems in machine learning. It is an ensemble learning method mostly used for correction of overfitting of the dataset. In this algorithm, we combine multiple classifiers to overcome a complex problem and in turn improve the accuracy and performance of the model. Random forest instead of relying on one decision tree takes majority votes of prediction and through that predicts the final output.

E.Support Vector Machine (SVM)

Support vector machine(or SVM) may be a supervised machine learning algorithm that's used both for classification and regression problems. The primary goal of SVM is to create a decision boundary that can segregate space into classes. The best decision boundary that is created by plotting each data point in n-dimensional space is called a hyperplane. SVM then selects the

extreme vectors which create the hyperplane, and classification is performed by finding the hyperplane that differentiates the two classes very well.

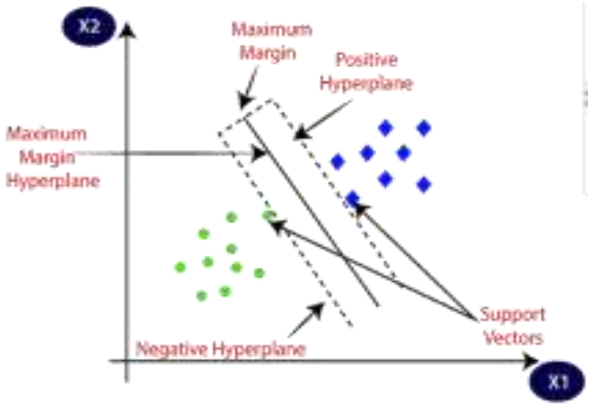


Fig. 3 SVM ALGORITHM

V. RESULTS

In the following table, TR is Training Accuracy and TE is Testing Accuracy, both of which are in percentage. The table shows the efficiency of various classification algorithms- heart disease, chronic kidney disease, and breast cancer.

MODEL	HEART DISEASE		KIDNEY DISEASE		BREAST CANCER	
	TR	TE	TR	TE	TR	TE
LOGISTIC REGRESSION	86.79	86.81	86.79	87.81	85.76	87.90
K-NEAREST NEIGHBORS CLASSIFIER	86.79	86.81	86.79	85.81	89.68	88.67
SUPPORT VECTOR MACHINE	93.40	87.91	93.40	95.20	88.56	86.70
DECISION TREE CLASSIFIER	100.00	78.20	89.05	77.20	93.45	91.23
RANDOM FOREST CLASSIFIER	100.00	87.20	91.37	89.20	94.37	96.25

TABLE 1: EFFICIENCY OF VARIOUS ALGORITHMS

Based on the total efficiency Random forest classifier was chosen for predicting Heart and Breast Cancer, however, the Support Vector Machine classifier (SVM) was chosen for predicting Chronic Kidney Diseases.

VI. CONCLUSION

Diagnosis of disease plays an essential role when it comes to the healthcare industry. The diagnosis of disease at an initial stage can help in a better treatment of an individual. This paper has implemented our ideas to benefit people who are not getting good health care and testing facilities. The proposed model uses a Decision tree classifier, random forest classifier, support vector machine(SVM) classifier, KNN, and logistic regression. The performance of the proposed model was compared with existing data mining models.

VII. FUTURE SCOPE

The proposed Health Diagnostic System represents an application that provides users to add their real-time data symptoms and know if they are suffering from the included diseases or not. This application can be expanded with the use of more advanced and efficient machine learning algorithms including deep learning which will also be able to take in symptoms through scanned images

like the X-Ray MRI scans, CT scans, ECG reports of the user and knows about various insights about the user's health condition. Diverse new diseases can be added that can be predicted using deep learning like lung cancer, pneumonia, etc. Deep learning constitutes a part of machine learning that helps us to increase the accuracy and prediction rate for our application and will make it more reliable with better results. It utilizes neural networks to increase the computational actions for better predictions. Therefore, the technology used has a bright future in the upcoming years and it can be of enormous help for the doctors in providing them with more rapid results.

VIII. ACKNOWLEDGEMENT

We express our deep gratitude to Mr. Amit Kumar Pandey, Assistant professor, Department of Information Technology for his valuable guidance and suggestion throughout my project work. We are thankful to Dr Preety Verma Dhaka, Project Coordinator for their valuable guidance. We would like to extend our sincere thanks to Prof. (Dr.) Prashant Singh, Head of Department, for his time to time suggestions to complete my project work. We are also thankful to Prof. (Dr.) Sanjay Kumar, Director for providing us the facilities to carry out the project work. We are thankful to and fortunate enough to get constant encouragement, support and guidance from all Teaching staff of the Information technology Department which helped us in successfully completing our project work. Also, we would like to extend our sincere esteems to all staff in the laboratory for their timely support.

REFERENCES

[1] K.Vembandasamy, et al Smart Health Care System using Data Mining, IJISET International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 9, September 2018-T [2] Awotunde et al. Medical Diagnosis System Using Fuzzy Logic 2014 Afr J Comp & ICT
[2] S. Kaur et al., "Medical Diagnostic Systems Using AI (AI) Algorithms: Principles and Perspectives," in IEEE Access, vol. 8, pp. 228049-228069, 2020, doi: 10.1109/ACCESS.2020.3042273
[3] M. S. van Mourik, A. Troelstra, W. W. van Solinge, K.G.Moons, and M.J.Bonten, "Automated surveillance for healthcare-associated infections: opportunities for improvement," Clinical infectious diseases, vol. 57, no. 1, pp. 85–93, 2013.
[4] P. Bhattacharjee, D. P. Edelson, and M.M.Churpek, "Identifying patients with sepsis on the hospital wards" Chest, vol. 151, no. 4, pp. 898–907, 2017.
[5] M. Sinha, J. Jupe, H. Mack, T. P. Coleman, S. M. Lawrence, and S. I.Fraley, "Emerging technologies for molecular diagnosis of sepsis," Clinical microbiology reviews, vol. 31, no. 2, 2018
[6] K. Gandhi, M. Mittal, N. Gupta, and S.Dhall, "Disease Prediction using MachineLearning," International Journal for Research in Engineering & Engineering Technology (IJRASET), vol. 8, no. 6, 2020.
[7] Nikita Pawar, et al. Decision Tree-based Health Prediction System (IJRASET) Volume 8 Issue VII July 2020.
[8] https://www.javatpoint.com/machine-learning-support-vector machine-algorithm

Exploring Security issues & Attacks in Cloud Computing Environment

Megha Gupta[1]
Assistant Professor,
Department of Computer Science & Engineering
Dr. Akhilesh Das Gupta Institute of Technology & Management, GGSIPU, New Delhi

***Abstract* - In today's world cloud computing is one of the most emerging technology because of its cost efficiency. Despite increased interest, the cloud computing is reaching the high pace. In this paper, I have characterize the security challenges and their impact on society. Also in this paper I have talked about equally important what is, the estimate of reliable computing and encryption support, with research advances in the cloud computing environment. There are different choices in world that are more common today, like business intelligence etc. Irrespective of what is company size or volume and magnitude of the cloud computing model, in this paper I have explained how technology virtualization could be used in responding to a Denial of Service attack (DoS).**

***Keywords*- Attacks, Cloud Computing environment, DoS, encryption, Security, decryption**

I. INTRODUCTION

The cloud computing environment means the delivery of scalable IT resources over the Internet. These resources can include applications and services, as well as the infrastructure on which they will operate. Cloud computing environment is a computing paradigm, where a large pool of systems are connected in private or public networks, to provide dynamically scalable infrastructure for different applications, data and file. With the advancement of this technology, the cost of computing, application hosting, content storage and delivery has reduced. When IT organizations manage a lot of data centers, the scope of the issues that are be addressed identified by the size of the data center. As IT organizations combine data centers, the problems to be addressed are getting bigger as well. One of the first things that many IT organizations will facing is that once you consolidate data centers and start building out a private cloud, your data centers has big security issues. In fact, those who has worked for an Internet service provider know how difficult and persistent those bad guys can be. And their most common form of attack is the distributed denial of service DDoS) [5,6].The motivation for these attacks used to be pretty undeveloped; they usually involve some form of force where the attacker asks for money to make the attacks disappear or some modern group simply wanted to make a political statement. They move data into the cloud provides great convenience to users. Cloud computing environment is a collection of all resources to enable resource sharing in terms of ascendable infrastructures, middleware and application development platforms, and value-added business applications [1] [2].

II. REQUIRED SECURITY IN CLOUD COMPUTING ENVIRONMENT

1- Securing data at relaxation

Cryptographic encryption techniques are the best practice in many U.S. states and countries worldwide, it's the law for securing data at the cloud provider end. Luckily, hard drive manufacturers are now shipping self-encrypting drives that implement the TCG's (Trusted Storage Group) standards. Self-encrypting drives build encryption hardware into the drive, providing automated encryption with minimal cost or performance impact. Software encryption can also be used, but it is slower and less secure since the encryption key can be copied off the machine without detection [3].

2- Securing data in transportation

Encryption techniques should also be used for data during transfer. In accumulation, authentication and integrity protection ensure that data only goes where the customer wants it to go and is not modified in transit. Well-established protocols such as SSL/TLS are used in this. The complicated part is strong authentication, as described next [3].

3- Verification

User verification is often the primary basis for access control, keeping the attacker out while allowing authorized users in with a minimum of worry. In the cloud environment access control are more important than ever since the cloud and all of its data are accessible to its permitted and regulatory experts inspect cloud provider policies and practices to ensure their adequacy. The issues to be considered include data security and export, compliance, auditing, data retention and destruction, and legal

discovery. In the areas of data retention and deletion, Trusted Storage and TPM access techniques can play a key role in limiting access to data.

4- Event response

As part of expecting the unexpected, customers need to plan for the possibility of cloud provider security breaches or user misbehavior. An automated response or at least automated notification is the best solution. TCG's IF-MAP (Metadata Access Protocol) specification enables the integration of different security systems and provides real-time notification of incidents and of user misbehavior.

III. SECURITY CHALLENGES OF CLOUD COMPUTING ENVIORNMENT

Regardless of its developing effect, concerns regarding cloud computing still present. In my opinion, the benefits outweigh the drawbacks and the computing is worth discovering. Some common challenges are:

1- Data Protection

Data Security is a crucial part that deserves scrutiny. Enterprises are unwilling to buy an assurance of business data security from vendors. They fear losing data to competition and the data confidentiality of consumers. In many instances, the actual storage location is not disclosed, adding onto the security concerns of enterprises. In the existing models, firewalls across data centers (owned by enterprises) protect this sensitive information. In the cloud model, Service providers are responsible for maintaining data security and business need to rely on them.

2- Recovery and Availability of data

All business applications have Service level agreements that are stringently followed. Operational teams play a key role in management of service level agreements and runtime governance of applications. In current environments, operational teams support appropriate clustering and fail over Data Replication System monitoring (Transactions monitoring, logs monitoring and others) Maintenance (Runtime Governance) Disaster recovery Capacity and performance management.
Management Capabilities Despite there being multiple cloud providers, the management of platform and infrastructure is still in its infancy. There is vast potential to improve on the scalability and load balancing features provided today.

3- Regulatory & Compliance Restrictions

In several countries, Government regulations do not allow client's personal information and other sensitive information to be physically located outside the state or country. In order to meet such requirements, cloud service providers need to setup a data center or a storage site exclusively within the country to comply with regulations. Having this type of organization may not always be possible and is a big challenge for cloud providers. With cloud computing environment, the action moves to the interface between service providers and multiple groups of service consumers. Cloud services will demand expertise in distributed services, procurement, risk assessment and service negotiation areas that many enterprises are only humbly equipped to handle.

IV. HOW to Handle Security Challenges in Cloud enviornment

The challenges in cloud computing environment are very alike to those of any other organization. Like IT companies, cloud providers have internal and external threats that can be moderated or accepted. Because of the commodity nature of IT, consumers expect the multi-tenancy aspect of cloud computing to bring their costs down [7].

Mini Cloud Risk/Mitigation Table	
Risk	Mitigation
Multi-tenancy	Infrastructure/data segregation
Ever developing risk	Continuing risk assessment program, CSO/CISO, Assessment
Relaxation of security	Periodic assessment/audit
Service provider tiers	Contract pass-through, coordinated security assessment
Contractor access	Background checks, Contracts, Segregation, Surveillance
Disasters	SLA, Multi facility provisioning
External Physical	Secure facility, Escort, Surveillance
External Logical	IPS, Firewalls, WAF, Secure Coding, Secure Architecture, Host hardening
Incidents	Facility & Per Customer Incident Response Plan
Application bugs	Layered security, Patching, Secure coding practices, Assessments, Segregation

DATA LEAKAGE	ENCRYPTION (AT REST & IN FLIGHT),

TABLE 1 MINI CLOUD RISK/MITIGATION TABLE

V. MALICIOUS ATTACKS IN CLOUD COMPUTING ENVIRONMENT

1- Denial-of-Service (DoS) Attack

DoS attack is a security attack [5] [6]. In this attack, an attacker attempts to make the resources of the victim devices unavailable to its intended users. Under this attack, an attacker launches SYN flooding, UDP flooding, and ICMP flooding attacks against the target node. An Attacker easily destroys the network or disables services provided by the target node without impersonating the source or destination node, modifying the data packets, or trying to decrypt encrypted data. An invader has to do is to send a bunch of data packets to the target node continually. These data packets occupy the network bandwidth and consume the target node's resources associated with various hardware elements. e.g. CPU and memory.

2- Distributed Denial-of-Service (DDoS)

DDoS is a kind of DoS attack [8]. The difference between DoS and DDoS attacks is that DDoS sends numerous malicious packets from multiple hosts to disable the services provided by the target node. DDoS attack generates more attack traffic than DoS.

3- Economic Denial of Service (EDoS)

EDoS attempts to consume IT resources maliciously that result in economic damage to their owners. Universities' resources in the cloud can be harmed by the following kinds of EDoS attacks [7]:

A. *Identity theft* – an invader hijacks the user accounts of universities' members and uses them for his personal gain or to damage universities economically.

B. *Resource Abuse* – If effective limits on the usage of paid resources from the cloud service providers, malicious actions can be made by attackers to create unexpected consumption of such resources.

C. *Public Channel Attack* – Cloud services delivered through public channel, such as metering per HTTPS requests, are vulnerable to attacks from the public Internet, such as Disturbed Denial of Service (DDoS). In the worst case scenario, EDoS eliminates the cost effective benefit of cloud computing and cause serious economic damage, even bankruptcy.

1.

4- Compromise of Service Engine

Cloud architecture relies on a highly specialized platform, the service engine that sits above the physical hardware resources and manages customer resources at different levels of abstraction. An attacker can compromise the service engine by hacking it from inside a virtual machine (IaaS clouds), the runtime environment (PaaS clouds), the application pool (SaaS clouds), or through its Application Programming Interface (API).

VI. PREVENTIONS OF ATTACKS

1- Intrusion Detection System

Intrusion detection technique has become an extremely feature of the system defense. In general, IDS collects network traffics, analyzes these traffics, and makes response or alerts the network to the manager if there is an intrusion taking place. Thus, the aim of the IDS is to alert or notify the system that some malicious activities have taken place and try to eliminate it [9]. According to the method of the collection of intrusion data, all the intrusion detection systems can be classified into two types: host-based and network-based IDSs. Host-based intrusion detection systems (HIDSs) analyze audit data collected by an operating system about the actions performed by users and applications; while network-based intrusion detection systems (NIDSs) analyze data collected from network packets. According to analysis techniques, IDS system is classified into two different parts: misuse detection and anomaly detection. Misuse detection systems use signature patterns of exited well-known attacks of the system to match and identify known intrusions. Anomaly detection systems identify those activities which deviate significantly from the established normal behaviors as anomalies. Anomaly detection systems tend to generate more false alarms than misuse detection systems because an anomaly may be a new normal behavior or an ordinary activity [8]. While IDS detect an intrusion attempt, IDS should report to the system administrator. There are three ways to report the detection results [4]: notification, manual response, and automatic response. In notification response system, IDS only generate reports and alerts. In manual response system, IDS provides additional capability for the system administrator to initiate a manual response. In automatic response system, IDS immediately respond to an intrusion through auto response system.

2- *Distributed IDS System*

Distributed IDS (DIDS) is a kind IDS designed to discover attacks on individual hosts as well as the network which connects them. A great deal of research has been devoted to providing reasonable solutions to support this framework [4][5][10][11]. The idea of DIDS is to aggregate data generated by individual intrusion detection systems. In DIDSs, messages sent by the agents are based on IDMEF [2] (Intrusion detection message format). It is an XML Document Type Definition for the exchange of messages between IDSs. The benefit of this kind of system is to gather the resources from IDSs in the network to withstand DoS or DDoS attack.

3- Using Intrusion Detection System in Virtual Machine

Intrusion Detection System is installed on the virtual switch which logs the network traffic in-bound and out-bound into the database for auditing. The packets are examined in real-time by the intrusion detection system for a particular type of attack based on predefined rules. The rules are defined based on well known attack strategies by the intruders. The IDS could determine the nature of attack and is capable of notifying virtual server the amount security risks involved [9].

VII. CONCLUSION

Security is a prime concern in every field so as in Cloud Computing environment. Attacks which can interrupt the security of Cloud Computing environment. Some of them are examined in this paper and techniques to handles such attacks are presented in different sections of paper. Additionally day by day new attacks are emerging and it is difficult to detect them. Therefore there is a requirement of prevention method which can detect and hold out all emerging attacks

REFERENCES

[1] B.R. Kandukuri, R. Paturi V, and A. Rakshit, "Cloud Security Issues," 2009 IEEE International Conference on Services Coming, Sep. 21-25, 2009, Bangalore, India, pp. 517-520.
[2] L.J. Zhang and Qun Zhou, "CCOA: Cloud Computing Open Architecture," 2009 IEEE International Conference on Web Services,July 6-10, 2009, Los Angeles, CA, USA, pp. 607-616
[3] Torry Harris, "cloud computing- An Overview" available: http://www.thbs.com/pdfs/Cloud-Computing-Overview.pdf,
[4] D.J. Ragsdale, C.A. Carver, Jr. J.W. Humphries, U.W. Pooch, "Adaptation techniques for intrusion detection and intrusion response systems," 2000 IEEE International Conference on Systems, Man, and Cybernetics, Vol.4 , 8-11 Oct. 2000 p.2344-p.2349.
[4] E.H, Spafford and D. Zamboni, "Intrusion Detection Using Autonomous Agent," Computer Networks, vol.34, issue 4, 2000,pp.547-570.
[5] G. Carl, G. Kesidis, R.R, Brooks, and S. Rai, "Denial-of-serive attack-detection techniques," IEEE Transaction on Internet Computing, Vol.10, issue 1, 2006, pp.82-89.
[6] J. Haggerty, S. Qi, and M. Merabti, "Early detection and prevention of denial-of-service attacks: a novel mechanism with propagated traced-back attack blocking," IEEE Journal on Selected Areas in Communications, Vol.23, Issue 10, Oct.2005, pp.1994-2002.
[7] Cloud Computing-Benefits, risks and recommendations for information security, ENISA, 2009, available: http://www.enisa.europa.eu/act/rm/files/deliverables/cloud-computing riskassessment/at_download/fullReport
[8] M.H., Islam, K. Nadeem, S.A., Khan, "Efficient placement of sensors for detection against distributed denial of service attack,"International Conference on Innovations in Information Technology,2008, 16-18 Dec. 2008, pp.653-657.
[9] AMAN BAKSHI, YOGESH B.,"Securing cloud from DDOS Attacks using Intrusion Detection System in Virtual Machine," 2010 Second International Conference on Communication Software and Networks,pp.260-264.
[10] S. Cheung, R. Crawford, and M. Dilger et al., "The Design of GrIDS:A Graph-Based Intrusion Detection System," Technical Report CSE-99-2, U.C. Davis Computer Science Department, January 1999.
[11] S.R. Snapp, J. Brentano, G.V. Dias, T.L. Goan, T. Grance, L.T.Heberlein, C.L. Ho, K.N. Levitt, B. Mukherjee, D.L. Mansur, K.L.Pon, and S.E. Smaha, " A system for distributed intrusion detection,"Compcon Spring'91, Feb-March 1991, pp.170-176.

A Green Solution To Cloud Computing Environment

Anjani Gupta
Dept. Of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology and Management
New Delhi, India
anjaniaggarwal.06@gmail.com

***Abstract*— Cloud computing is a dynamic field of information and communication technologies (ICTs), who brought the revolution of modern networking and also introduces new challenges for environmental protection. In recent time, computing technologies and thoughts are shifted to remote data centers and the software and hardware services available on the basis of pay for use. However, increase in the number of data centers, hosting cloud computing applications consume huge amounts of energy, contributing to high costs and carbon footprints to the environment. To overcome this problem, Green Computing came into subsistence. Green computing is a research topic to address climate and energy challenges. Green Cloud computing is visioned to achieve not only efficient processing and utilization of computing environment, but also minimize the energy consumption. This paper aims to discuss the potential of green cloud computing, its future trends and measures to improve the efficiency of data centers which will in turn reduce the emission of CO_2. This generates research challenges when such energy-saving technologies are required to minimize the bad impact of Cloud computing on the environment.**

***Keywords*— *Cloud Computing, Energy-efficiency, Efficient Data Centres, Co_2 emissions*.**

I. INTRODUCTION

The evolution of cloud computing which provides on-demand provisioning of elastic resources with pay-as-you-go model has transformed the Information and Communication Technology (ICT) industry. Over the last few years, large enterprises and government organizations have migrated their data and mission-critical workloads into the cloud. As we are moving towards the fifth generation of cellular communication systems (5G), Mobile Network Operators (MNO) need to address the increasing demand for more bandwidth and critical latency applications [10]. Thus, they leverage the capabilities of cloud computing and run their network elements into distributed cloud resources.

The adoption of cloud computing by many industries has resulted in the establishment of humongous data centers around the world containing thousands of servers and network equipment. In the last few years, Cloud support data centers are increasing greatly because of the demand for processor resource. Since more data centers are came into survival the energy utilization of these data centers are also increased to a great extent [1]. In adding to high energy utilization there is an adding impact on the surroundings by the form of carbon dioxide emission.

Fig. 1. Cloud Computing

Green cloud computing found its root when the National Institute of Standards and Technology (NIST) made a point on drawback of cloud computing. According to NIST- “the major objective of cloud computing is to provide the shared resources and at the same time the disadvantage is its high infrastructure cost and the unnecessary power consumption and carbon emission

as shown in Fig.2". As cloud computing becomes more common and demands on the internet grow, major companies hosting online services are using more and more energy for their data centers [2].

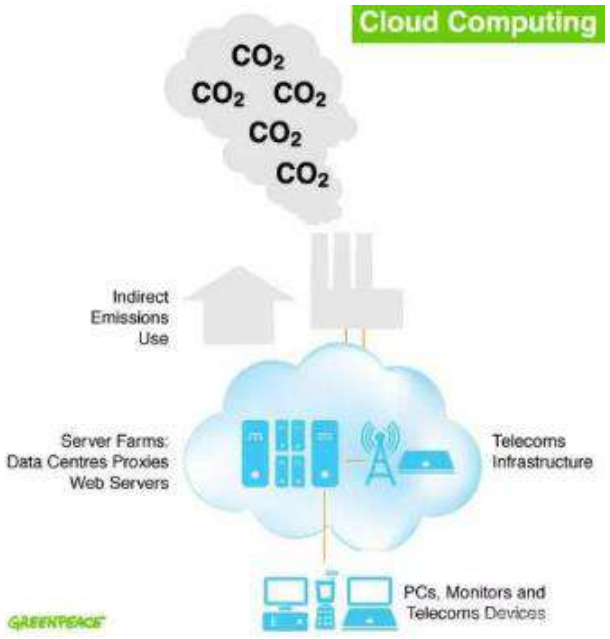


Fig. 2. Cloud Computing and its effect on climate change

To address this problem, data center resources need to be managed in an energy efficient manner to drive Green Cloud Computing. In particular, Cloud resources need to be assigned not only to satisfy the quality of service requirements specified by users, but also to reduce energy usage.

II. NEED FOR GREEN CLOUD COMPUTING

Computing needs to be "green," meaning that we are conscious of the ways our lives and our work impact the world. Green computing, also called green technology, it is environmentally responsible and eco-friendly use of computers and their resources. This involves the implementation of energy-efficient central processing units (CPU"s), servers and peripherals as well as reduced resource consumption and proper disposal of electronic waste (e-waste). It is also the development of environment friendly products, energy efficient computers and recycling the system products. Increased power use has increased the IT costs [3].

Growing risks on climate change and desire to lessen energy use has given rise to green computing. Main goals of green computing are to reduce the use of toxic and hazards materials and improve the energy efficiency, recycling of e-waste. Such practice includes the efficient implementation of server and other subsystems as well as reduce in the power consumption. [4]. Research continues into key areas such as making the use of computers as energy efficient as possible, and designing algorithms and systems for efficiency related computer technologies.

III. APROACH TO GREEN CLOUD COMPUTING

The important technology for energy efficient in clouds is the "Virtualization" Which helps in significant improvement in energy efficient by sharing the same infrastructure. By using virtualization companies can gain high savings in the form of space, management, and energy.

Dynamic Provisioning

In traditional computing, data centres and private infrastructure used to be maintained to fulfil worst case demand. Thus most of the IT companies deploy more infrastructure than needed. The reasons for this situation are[6]:

- It is very difficult to predict the demand before itself.
- To guarantee availability of services and to maintain certain level of service quality to end users.

Multi-tenancy

Cloud computing infrastructure reduces overall energy usage and associated carbon emissions. The SaaS providers serve multiple companies on same infrastructure and software. This approach is more energy efficient than multiple copies of software

installed on different infrastructure, which can minimize the need for extra infrastructure. The smaller fluctuation in demand results in better prediction and results in greater energy savings.

Server Utilization

Using virtualization technologies, multiple applications can be hosted and executed on the same server in isolation, thus lead to utilization levels up to 70%[7]. Even though high utilization of servers results in more power consumption, server running at higher utilization can process more workload with similar power usage.

Data Centre Efficiency

By using the most energy efficient technologies, Cloud providers can significantly improve the Power Usage Effectiveness (PUE) of their data centres. Cloud computing allows services to be moved between multiple data centre which are running with better PUE values.

IV. DATA CENTRE ENERGY CONSUMPTION

The architecture of a data centre is complex since it does not only consist of the hardware elements but also the software that runs in the IT infrastructure. Therefore, we can categorize its elements into two layers which are hardware and software, as shown in Fig. 3.

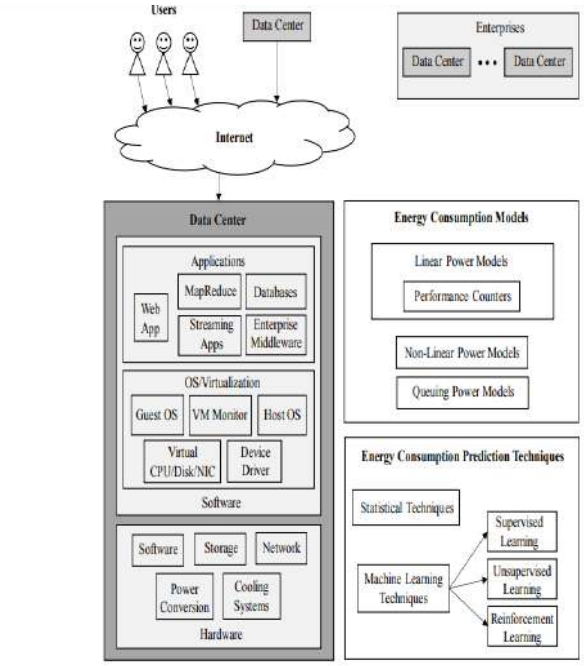


Fig. 3. Categorization of Data Centers: Hardware and Software

Data centers typically are powered by electricity. However, following the strategy for decreasing carbon emissions and complying with sustainable operational models, modern data centers use alternative energy sources such as geothermal, wind and solar power. The electric power flows from external power grids into internal infrastructure facilities, Information Technology (IT) equipment and other support systems. The energy flows to the internal IT facilities through Uninterrupted Power Supplies (UPS) to maintain a consistent power distribution even during possible power failures[8].

V. MEASURES TOWARD GREEN CLOUD COMPUTING

There are a hardware and software solutions that can be taken to significantly curb the power consumption and impact on environment. The key techniques to reduce the energy consumption, CO_2 emission and progress towards green data centers are [11]:

Application Oriented

SaaS model has changed the way applications and software are distributed and used. More and more companies are switching to SaaS Clouds to minimize their IT cost. Thus, it has become very important to address the energy efficiency at application level itself.

Virtualization

It is an energy efficient method where we can share the available physical resources among virtual machines (VMs). Dynamic resource allocation is a technique which allow us to utilize and assign dynamically free computing resources among VMs, saving at the same time considerable amount of energy[13]. Virtualization is an efficient way to provide server consolidation and power off the server that operate in idle mode. In many cases, an idle server consumes 70% of the power consumed by a server running at the full CPU load [9].

Energy Efficient Hardware

Computer systems are made up of hardware i.e. processor onboard graphics, disk, fan, etc. these hardware should consume less power. Computer power can be saved by means of various well-known techniques[14]. Firstly, the system should be shut down by sensing lack of interaction. Secondly by slowing down the CPU clock speed known as clock gating or we can power off the chips also which is known as chip gating. Other methods of powered down the system processor are Speed Step, Power Now, Cool'nQuiet [15] or Demand-Based Switching.

Monitoring/Metering

It is said that you cannot improve what you do not measure. It is essential to construct power models that allow the system to know the energy consumed by a particular device, and how it can be reduced. To measure the unified efficiency of a datacenter and improve its' performance per-watt, the Green Grid has proposed two specific metrics known as the Power Usage Effectiveness (PUE) and Datacenter Infrastructure Efficiency (DciE)[12].

PUE = Total Facility Power/IT Equipment Power

DciE = 1/PUE = IT Equipment Power/Total Facility

Power x 100%

Network Infrastructure

The energy efficiency issues in networking is usually referred to as "green networking", which relates to embedding energy-awareness in the design, in the devices and in the protocols of networks.

VI. CONCLUSION & FUTURE SCOPE

This paper gives a brief study about green computing in a cloud environment. Green cloud computing is visioned to achieve not only efficient processing and utilization of computing environment, but also minimize the energy consumption. The management of power consumption in data centres has led to a number of substantial improvements in energy efficiency. In conclusion, we can say that no technology or method is green but its user who uses it makes him green by using it effectively according to its requirement. Cloud Providers need to minimize the consumption of its electricity demand and take major steps in using renewable energy sources rather than just looking for cost minimization However, green computing technology needs to be further researched for a better and efficient use of cloud.

REFERENCES

[1] Barham, P., Dragovic, B., Fraser, K., Hand, S., Harris, T., Ho, A., Neugebauer,R., Pratt and I. Warfield "The art of virtualization," in ACM IGOPS, Operating Systems Review, 37 (5), pp. 164–177.

[2] Thanmai.C , Dr.K.N.Narsimaha Murthy and Dr S. Ambareesh "A Survey on Evolution of Green Computing In the Cloud Environment," in International Journal of Innovative Research in Computer and Communication Engineering, Vol.3, Special Issue 5, May 2015.

[3] Biswajit Saha,"Green Computing," in International Journal of Computer Trends and Technology (IJCTT) ,Volume 14, No.2, Aug 2014.

[4] Archana Tahiliani and Apexa Purohit," A Review on Energy Saving Using Green Computing System", International Journal of Engineering Development and Research, ISSN: 2321-9939.

[5] L. Krug, M. Shackleton, and F. Saffre, "Understanding the environmental costs of fixed line networking," in Proc. 5th Int. Conf. Future e-Energy Syst., 2014, pp. 87–95.

[6] Archana Tahiliani, Apexa Purohit," A Review on Energy Saving Using Green Computing System" International Journal of Engineering Development and Research, ISSN: 2321-9939.

[7] Ankita Atrey, Nikita Jain and Iyengar N.Ch.S.N,"A Study on Green Cloud Computing", International Journal of Grid and Distributed Computing Vol.6, No.6 (2013), pp.93-102.

[8] Gaurav Jindal.; Manisha Gupta,"Green Computing "Future of Computers", International Journal of Emerging Research in Management &Technology, December 2012.

[9] A. Beloglazov, J. Abawajy and R. Buyya, "Energy-Aware Resource Allocation Heuristics for Efficient Management of Data Centers for Cloud Computing", Article in Future Generation Computer Systems, May 2012.

[10] K. Bilal, S. Malik, S. Khan, and A. Zomaya, "Trends and challenges in cloud datacenters," IEEE Cloud Comput., vol. 1, no. 1, pp. 10–20, May 2014.

[11] B. Whitehead, D. Andrews, A. Shah, and G. Maidment, "Assessing the environmental impact of data centres—Part 1: Background, energy use and metrics," Building Environ., vol. 82, pp. 151–159, Dec. 2014.

[12] P. Corcoran and A. Andrae, "Emerging trends in electricity consumption for consumer ICT," Nat. Univ. Ireland, Galway, Ireland, Tech. Rep., 2013.

[13] H. Liu, C.-Z. Xu, H. Jin, J. Gong, and X. Liao, "Performance and energy modeling for live migration of virtual machines," in Proc. 20th Int. Symp. HPDC, 2011, pp. 171–182.

[14] A. Beloglazov and R. Buyya, "Energy efficient resource management in virtualized cloud data centers," in Proc. 10th IEEE/ACM Int. CCGrid, May 2010, pp. 826–831.

[15] E. Feller, C. Rohr, D. Margery, and C. Morin, "Energy management in iaas clouds: A holistic approach," in Proc. IEEE 5th Int. Conf.

Autonomous Self Driving Cars using CNN

[1]Aashita Chhabra, [2]Gursahaj Bir Singh Kohli
[1,2]Assistant Professor
[1,2]Dr. Akhilesh Das Gupta Institute of Technology and Management
New Delhi, India
[1]aashita.chhabra@gmail.com
[2]sahaj939@gmail.com

***Abstract*- Computer vision is an integral part of today's scientific research, that deals with how computers can be made to gain high-level understanding from digital images or videos. This aims to train the computer to carry out tasks in a similar way that the human visual system does. By training the computer systems to recognize objects, we can further implement this to carry out actions depending on the objects detected. This approach can be used in making Autonomous cars. Autonomous cars are the one that run on it's own without any human intervention. This means that the car would be able to analyze all the objects on the road, and take appropriate actions depending on what it sees. The car hence made, is intelligent enough so that it follows lane driving at an appropriate pace so as to avoid mishappenings and move smoothly on the road. By using convolutional neural networks, the car is able to make relevant decisions in impromptu situations. This paper will help minimizing mishaps by eliminating human errors thereby helping mankind. Since this car is based on a hands-free driving approach, it will be a boon for differently abled persons.**

Keywords - Autonomous car, computer vision, vision, neural networks, Keras.

I. INTRODUCTION

From the past few years, as one of the most engrossing technology trends in the automotive industry, autonomous vehicles have received increasingly significant attention due to their significant potential in enhancing vehicle safety and performance, traffic efficiency, energy saving and their ability to run without any human intervention thereby being a support for differently abled persons. Many companies are actively developing related hardware and software technologies towards the cars, that can possess self driving capabilities. Deep neural networks (DNNs) have been successfully applied in various perception and control tasks in recent years.

There are various challenges that the car must cope with in order to run successfully on the road. First, computations made by the cars should be done in real time with high accuracy. For example, calculating the ROI for various objects on the road, that is directly linked to the safety of the vehicle. For these calculations, high computational power is required that must deliver the precise results in time. Whereas, the high computing hardware must also keep in mind about the cost, size and power of the hardware.

To get an idea of how the vehicle would operate, we need testing environment and workload, but practicing the project in real environments would not only be expensive, but can cause hindrance in our project because of safety issues. So, in order to understand the working, we can use open-source emulators that depict how the car can run autonomously.

In this paper, we present how autonomous vehicles possess deep learning in real time, that is based on deep convolutional neural network (CNN). In a real-time environment, CNN makes use of a camera, as a receptor, that inputs data in video format. The video is processed by converting each frame into an image, which is then used to analyze the environment from that image, so as to predict the steering angles and speed, in order to move the car smoothly. Depending on the type of network, a CNN can have various layers, which has numerous connections and weights associated with it.

Our main observations are as follows. First, we will find out how CNN interfacing is feasible in computing analysis on videos, that helps in autonomous movement of vehicles. We will be discussing it with reference to the Nvidia end-to-end Deep Learning CNN model. Secondly, we will suggest some algorithms, other than CNN that perform better, with the resources allocated.

The rest of the paper is organized as follows. In Section II, we review the system architecture of Tesla's self-driving car. In Section III, we study the background deep convolutional neural networks and training methods. In Section IV, we study the actual implementation that must be opted in order to run the car autonomously. Section V concludes the paper with brief discussion. Section VI describes the future scope that can be adopted in order to enhance the autonomous driving experience.

II. SYSTEM ARCHITECTURE OF TESLA'S SELF-DRIVINGCAR

The autopilot mode in Tesla's cars has made it possible for the drivers to run the car without human intervention. These cars have 8 surround cameras, that provides 360° visibility around the car at upto 250 meters of range.

It has twelve updated ultrasonic sensors that complement its vision, allowing for detection of both hard and soft objects at

nearly twice the distance of the vehicle. A forward-facing radar with enhanced processing provides additional data about the world on a redundant wavelength that is able to see through heavy rain, fog, dust and even the car ahead. Keeping in mind the high computation, Tesla cars have an onboard computer with 40 times the computing powerthan the previous generations, to run the Tesla neural network for radar, sonar and vision processing.

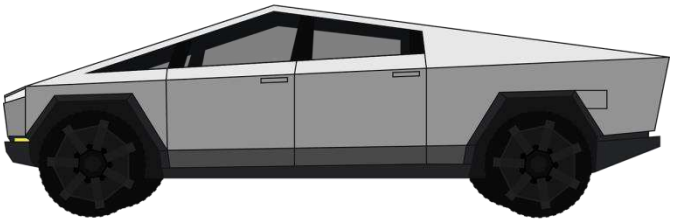


Fig. 1 Tesla Cybertruck

All new Tesla cars have the hardware needed in the future for full self-driving in almost all circumstances. The system is designed to be able to conduct short and long-distance trips with no action required by the person in the driver's seat.

These cars are programmed in such a manner, that they can locate parking space on their own, and would come back to the rider, when he calls the car through his mobile device.

III. BACKGROUND

In this section, we would see how classification and training methods are opted for training the vehicle to run autonomously.

A. Deep convolutional neuralnetwork: Deep convolutional neural network is based on the standard Convolutional Neural Network (CNN) just by increasing the number of layers, in order to make it more susceptible for learning. Its main components are:

1) Convolution: A series of convolution layers are connected to one another, such that output of one becomes the feeder to the next layer. The difference between the layers depends on the weights associated with each of the layers. When the data is input in these layers, filters are used that depict the characteristics of the image. So, each filter is used as a feature detector to help classify the image. Hence, applying convolution over an image helps in extraction of feature maps from an image.
2) Max Pooling: The feature maps that are generated from convolutional layers are sub sampled in order to gain specific information. Max pooling is used to reduce the dimensionality of feature maps, by retaining only the essential features that are consistent with our feature map. Because the dimensions are reduced, without affecting the information, the computational complexity is also reduced.
3) Fully connected layer: The output from the previous layers, is flattened and fed into the fully connected layer where the number of pixels correspond to input node of this layer. These fully connected layers are responsible for taking the features as inputs, processing them to find the probabilities in order to obtain the class of what the image belongs to.

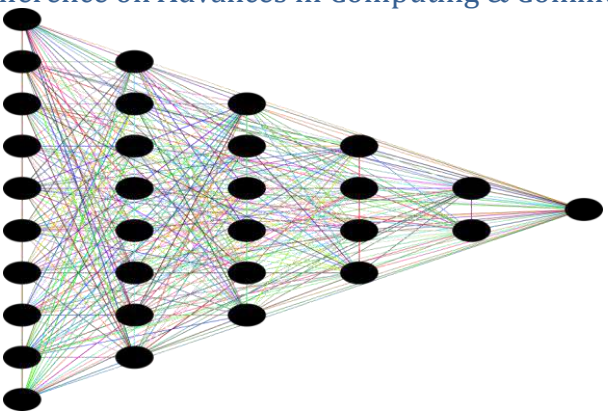


Fig. 2 Representation of fully connected layers

4) Activation function: Activation function is used to define the output of a node. Depending on the input values, activation function is used to predict the output values. It is used to transform non-linearity to linear data received from fully connected layers. Some examples of Activation functions are hyperbolic tangent function, sigmoid function and rectified linear units.

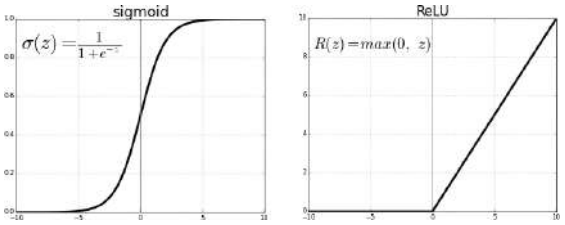


Fig. 3 Sigmoid and Relu functions

Convolution layers and max pooling layers are laid successfully, followed by fully connected layers, in order to create a deep convolutional neural network for classification of images.

B. *TrainingMethods:* The training of data can be done by using gradient descent method, but the problem of computational complexity arises, as training of entire batch in one go is not a feasible solution. So we make use of small batches of the training data, and then train them accordingly. We make use of epochs, so that if the result gets trapped in local maxima, then its jumped out to get optimized results. As gradient descent algorithm works by updating the weights and biases such that error function gets decreased, we need to make sure that the entire training data is run more than once. So more than one epochs must be used, but not too many as it may lead to overfitting of the data.

The cost function is calculated as:

As we know that equation of line is

$$y = mx + b$$

where,

y: Y coordinate of the line x: X coordinate of the line m: Gradient of the line

b: Y intercept of the line

So, initially, the model will find the value of 'y' based on the weights and biases that it has. Then, based on the training data, it will compare the output and calculate the value from error function.

The error function is calculated as:

$$Co=\frac{1}{N}\sum(Yi'-Yi)^2$$

computationally exhaustive than real time computation because, the model is to be trained on the test dataset, over multiple epochs, for higher optimality, whereas in real time computation, inferencing is done on the basis of model generated.

In the training phase, the data that is fed should be such that the model must not result in over or under fitting. Adequate data must be fed in order to perfectly train the model. Hence for that, we must divide our entire training dataset into test set and validation set. The test set would be the one on which the model would be trained, whereas, the validation set would be which the model has never seen before. As our model is encountered with the data it has never seen, we will calculate its loss, called as Validation loss, and compare it with the loss we get from test dataset. If the validation loss is much lower than training loss, then this is a case of underfitting, whereas if the validation loss is more, then it arises overfitting of data.

For training of the model, the NVIDIA's DAVE-2's convolutional neural network (CNN) can be inferred. This model takes raw image (200 x 66 RGB pixels) as input and produces steering angle as output. The figure shows the network architecture used in this paper.

where,i=1

y_i' = actual value from training dataset

y_i = value calculated using weights and biases

Depending on the value from cost function, back propagation is done, that changes the values of weights and biases, in order to get optimum results. This gradient descent algorithm works in such a manner, that when it is away from the optimum result, it makes greater leaps, but as soon as it comes near to the result, its leaps reduce, so that the optima is not skipped. The size of steps taken to reach is termed as Learning Rate. So initially, the learning rate is higher, but as it tends to reach the optimum result, the learning rate tunes itself and reduces in size which prevents the overshooting of minima.

IV.IMPLEMENTATION

Autonomous driving can be implemented, by dividing the problem into subproblems, and working on each subproblem in a concomitant manner such that the result of all subproblems lead to the optimum result of that problem at that instant. For instance, the problems can be divided as, lane driving, car ROI calculation, sign detection, all must be solved at every instant of driving.

The implementation is based on two phases, first being training of data, which uses backpropagation, to update their weights to minimize the error and second phase is inferencing in which, unseen data is fed to the system to see if correct computations can be made. The training phase is more

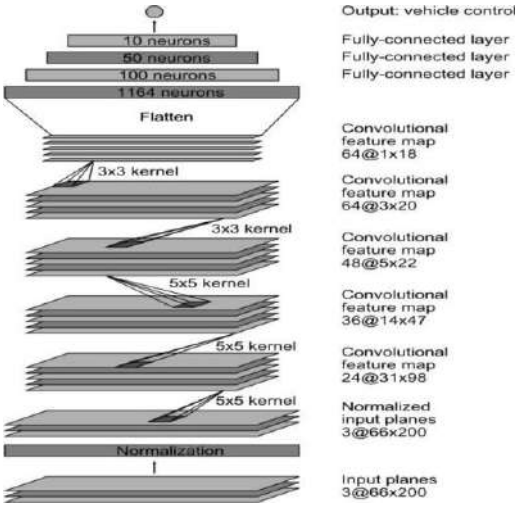


Fig. 4 NVIDIA CNN Model

This network comprises 9 layers, 250K parameters and 27 million connections. Initially, the raw image is fed to the network, to which normalisation is done, in order to analyze only the relevant part of the image. Then the normalised image with 3 RGB channels and 200 x 66 px, is passed to the first convolutional layer in which a 5 x 5 kernel is used for feature extraction. Then further, a series of 4 convolutional layers are used with varying kernel sizes, in order to extract required featuresfrom the image. Then, the output from convolutional layers is flattened, in order to serve as an input for fully connected layers. 4 fully connected layers, with varying neurons, run together, in order to depict the steering angle that can be inferred from the image.

The code for the following model in Python programming language would be as follows:

```
defnvidia_model(): model = Sequential()
    model.add(Convolution2D(24, 5, 5, subsample = (2, 2),
input_shape = (66, 200, 3), activation= 'elu' ))
   model.add(Convolution2D(36, 5, 5, subsample = (2,2), activation ='elu'))
   model.add(Convolution2D(48, 5, 5, subsample = (2,2), activation ='elu'))
    model.add(Convolution2D(64, 3, 3, activation = 'elu'))
    model.add(Convolution2D(64, 3, 3, activation = 'elu') model.add(Flatten())
    model.add(Dense(1164, activation = 'elu')) model.add(Dense(100, activation = 'elu')) model.add(Dense(50, activation =
    'elu')) model.add(Dense(10, activation = 'elu')) model.add(Dense(1))
```

The generated model can then be trained by feeding it up with test data and comparing it with validation data, in order to see the fitting of the curve.

A. *Collection of data:* The data, on which the model is trained, must be collected from the environment, where the vehicle would run in future. Many open source emulators are available online that can be used in recording data. These emulators help in depicting the behaviour the cars would possess in the real world, when they encounter similar situations. As most of the driving is done along the straight path, the data collected would be biased towards travelling straight, rather than making turns, that would topple the car. So correct analysis of data must be done in order to remove any kind of biasness.

B.

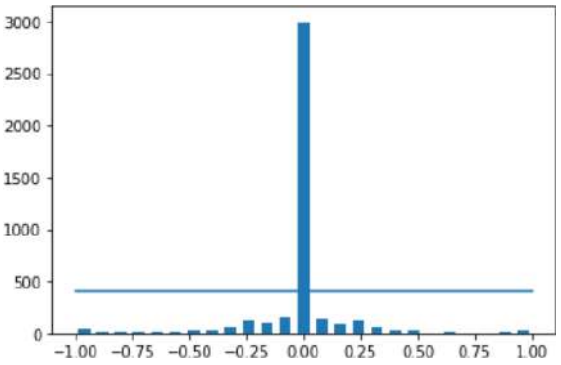


Fig. 5 Graph representing frequency to steering angles

The Fig shows the frequencies in Y axis and steering angle in X axis. Here, we can clearly see that the frequency with respect to steering angle 0 has a much larger value than left and right steering angles. This needs to be controlled to a comparable value with left and right steering angles, because of which we select an arbitrary value, like here we can see that the value selected is 400. Now, as the frequency gets comparably distributed, the biasness is reduced at a much larger extent thereby, preventing the car from moving in a particular direction.

C. Data augmentation: Data collection is done in by recording the values when the car is run in the domain, but that data is not sufficient, as what the receptor actually perceives, may be different than what it is trained for. For instance, let us suppose that the car is supposed to take a left turn, but according to the data for which it is trained, there is difference in visual due to environmental factors, like either it might be gloomy, or cloudy. So now the car is unaware of its next action to be taken, as it was never trained for such conditions. So, for this, we use the technique of Data Augmentation.

Data augmentation is the means of selecting random data, from the dataset, and introducing various features to the image, so as to increase the size of the dataset, in order to train the car to maximum kinds of possibilities. Some data augmentation techniques are, flipping the image vertically, zooming the image, changing the brightness and translation of image on X or Y-axis.

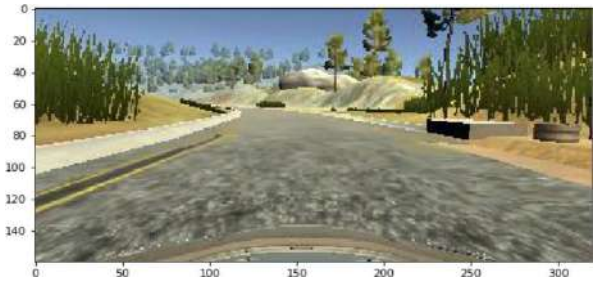


Fig. 6 Original image of what the receptor captures

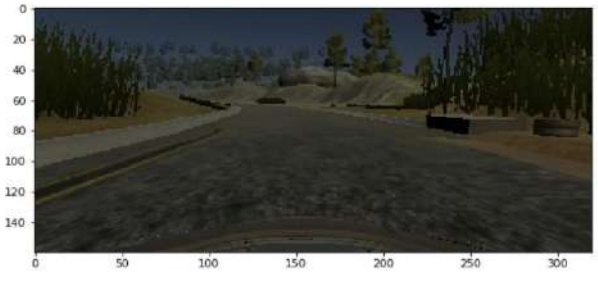


Fig.7 Reduced brightness of image

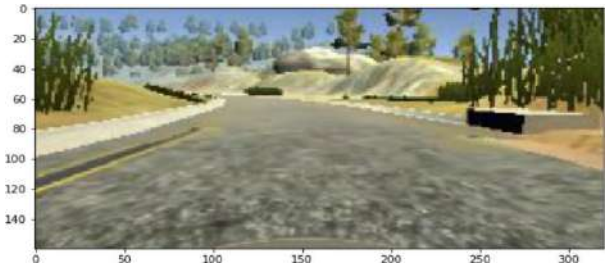


Fig. 8 Zoomed image

Fig. 9 Transitional image from both axes

The above figures show how the original images can be augmented in different ways, to train the car for all such experiences, so that it is capable of handling maximum situations.

The essential feature of using data augmentation technique is the fact that augmentation is done 'over the fly'. This means that augmented images are not added directly to the database, rather they get generated as and when required. Moreover, the images generated do not occupy much space, thereby reducing computational complexity.

After data augmentation, the car has a better understanding of what steering angle is to be selected, as more data gets generated from existing data, car knows what angle is to be taken at what condition.

V. CONCLUSION

In this paper, we presented a deep learning approach, that can be used for autonomous driving of vehicles. We explained the system architecture of Tesla's Cybertruck. Then we gave a background of deep convolutional network and the training methods used. Then, using CNN, feature extraction was performed on images that the car receives, so as to find out the steering angles, that need to be opted, in order to move the car in that particular direction autonomously. We also inferred the NVIDIA's end to end deep learning model, that can be implemented in order to predict the steering angles. Finally, we implemented the data augmentation strategy, that was used to increase the dataset, by various methods, in order to train the car in various conditions so that the car is capable of making decisions in different conditions. These methods can be used to train the car to move autonomously without any human

VI. FUTURE ASPECTS

The paper we presented is based on convolutional neural network, for feature extraction and analysis of data. However, using other methods, we can have better speed and higher accuracy.
Using Single Shot Detection (SSD), You Only Look Once (YOLO) and Faster RCNN can help in achieving better results.

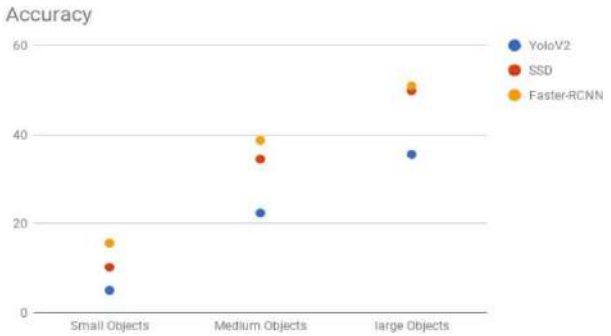


Fig. 10 Accuracy of different methods

Depending on what type of objects are to be tracked, various methods can be opted. Figure 10 depicts the level of accuracy that can be achieved for varying size objects.

REFERENCES

[1] ShahrzadFaghih-Roohi∗,SiamakHajizadeh†,AlfredoNu´nez˜†,RobertBabuska∗and Bart De Schutter∗, "Deep Convolutional Neural Networks forDetectionofRailSurfaceDefects"

[2] Hongbo Gao, Bo Chen, JianqiangWang, Keqiang Li, Jianhui Zhao, and Deyi Li , "Object Classification using CNN-Based Fusion of Vision and LIDAR in Autonomous VehicleEnvironment

[3] JieTang1 , Yong Ren2 , Shaoshan Liu2 , "Real-Time Robot Localization, Vision, and Speech Recognition on Nvidia JetsonTX1"

[4] Michael G. Bechtel†, Elise McEllhiney†, Minje Kim, HeechulYun†, "DeepPicar: A Low-cost Deep Neural Network-based AutonomousCar"

[5] Goodfellow, I.J., Bulatov, Y., Ibarz, J., Arnoud, S., Shet, V.: Multi-digit number recognition from street view imagery using deep convolutional neural networks. arXiv preprint arXiv:1312.6082(2013).

[6] Escalera, S., Bar, X., Gonzlez, J., Bautista, M.A., Madadi, M., Reyes, M., Ponce, V., Escalante, H.J., Shotton, J., Guyon, I.: Chalearn looking at people challenge 2014: Dataset and results. In: ECCV workshop(2014).

[7] Poppe, R.: A survey on vision-based human action recognition. Image and vision computing 28(6), 976–990 (2010).

[8] J. Nagi, F. Ducatelle, G. Di Caro, D. Ciresan, U.Meier,

[9] Giusti, F. Nagi, J. Schmidhuber, and L. Gambardella. Max-pooling convolutional neural networks for vision-based hand gesture recognition. In Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference on, pages 342–347, Nov 2011.

[10] Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R.: Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580(2012).

[11] Glorot, X., Bordes, A., Bengio, Y.: Deep Sparse Rectifier Networks. Proceedings of the 14th International Con- ference on Artificial Intelligence and Statistics 15, 315–323(2011), http://eprints.pascal-network.org/archive/ 00008596/

[12] JeongwooJu, Heechul Jung, Junmo Kim: Image Classification Using Convolutional Neural Networks With Multi-stageFeature.

[13] Mark Dougherty:Are view of neural networks applied to transport

A Brief Analysis of Blockchain Technology

Arushi Gupta[1], Gaurav Kumar[2], Nishtha Kapoor[3], Saijal Gupta[4]

[#]*Department of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology & Management New Delhi*

[1]arushigupta568@gmail.com

[2]laxaman.pal@gmail.com

[3]nishthakapoor2017@gmail.com

[4]saijalgupta92@gmail.com

***Abstract*— Technology has affected every part of our lives and that also includes our financial systems. Terms like cryptocurrencies and Bitcoin are more popular than ever and Blockchain Technology is one of it's basic foundations. Block chain, in simple terms, is a decentralized and distributed ledger system that ensures the integrity and immutability of document by recording the provenance of a digital asset. In this paper we discuss both basic and advanced concepts regarding the working and functioning of the Blockchain technology and the challenges it faces in the current scenario. We also discuss the future prospects of blockchain technology in financial services, Internet of Things, Trade Processing, personal identification and so on.**

***Keyword*— Decentralized, ledger system, block chain, immutable, scalability.**

I. INTRODUCTION

Since and after it's tremendous reach of $19,783.06, Bitcoin [1] has been the talk of the wall street even after more than a decade. During this whole journey, it has seen its highs and lows and it still is one of the most widespread and popular cryptocurrency ever created. Now, if something is this popular, there has got to be something special about the mechanism it's working upon. The technology Bitcoin works upon, Blockchain technology, introduced in the year 2009, is simply a decentralized and distributed ledger system that records the provenance (place of origin) of a digital asset.

It makes the history of any digital asset unalterable and transparent through the use of concepts like decentralization and cryptographic hashing [2]. In contrast to a centralized system, a decentralized system provides every node present on the network with an equal status and responsibilities.

The goal of a blockchain is to let people share assets in a secure and tamper proof way. Since it is a decentralized system, a single party won't have authority over the whole chain. Every individual device over the network called nodes has their own copy of blockchain and can update, verify and validate the record independently. This brings transparency to the whole network. As each transaction occurs it is put into a block. Each block is connected to the one before and after itself. Groups of transactions are blocked together and fingerprint, i.e., hash of each block is added to the next thus creating an irreversible chain. It makes it impossible to alter the blockchain thus making the system immutable and enhance trust and integrity of the system. Blockchain also eradicates the use of an intermediary in any kind of operation, thus it has been proven to be a vital technology in the field of financial transactions and literally any other field that requires a ledger system.

The rest of this paper is organized as follows. Section II shows the characteristics of blockchain technology. Section III summarizes the architecture of blockchain. Section IV gives the classification of blockchain technology. Section V explains the working of a blockchain. Section VI discusses the various consensus models and algorithms in the industry. Section VII introduces the challenges faced by blockchain technology. Section VIII summarizes the applications to blockchain in the various sectors and countries. Section IX discusses the future aspects and trends possible in the blockchain technology and Section X concludes the paper.

II. CHARACTERISTICS OF BLOCKCHAIN

The three main pillars of Blockchain Technology [3] which have helped it gain widespread acclaim are as follows:

- Decentralization
- Transparency
- Immutability

A. Decentralization:

In a centralized system all the intermediation, i.e., one or few trusted parties carry out most of the intermediation tasks for vast network of users, thus limiting all the power in the hands of central party only. In contrast, in a decentralized system all the nodes in the network carry out the tasks. There is no need for a third party to maintain data consistency.

B. Transparency:

Blockchain technology is nothing but sequence of blocks and each block is linked to the previous and next block. It records all sequences of transaction from beginning to end thus recording the provenance of each transaction. A transparent ledger of changes preserves integrity of the document, which creates trust in the asset.

C. Immutability:

Immutability [2], in the context of the blockchain, means that once something has been entered into the blockchain, it cannot be tampered with.

Blockchain uses cryptographic hash function to achieve immutability in hashing algorithm, the system takes a piece of data of any length and creates a unique fingerprint of a small fixed length hashing function almost uniquely compresses a data input. Bitcoin uses SHA-256(Secure Hashing Algorithm 256). A good cryptographic hash function is concealing, deterministic, and has a high "Avalanche effect", thus changing a tiny bit of input can alter the hash dramatically.

III. ARCHITECTURE OF BLOCKCHAIN

A blockchain can be defined as a family of decentralized record keeping systems in which each transaction created is distributed over the network as shown in fig. 1. Blockchain is a sequence of blocks, which holds a complete list of transaction records like conventional public ledger. Blockchain technology is based on three main concepts which are explained below:

A. Blocks:

To conserve computing power, most blockchains use the concept of "Batch processing". Nodes group pending transaction into blocks [3]. A block consists of the data, i.e., transaction. In particular, a block is divided into two parts- Block header and Block body. Block body comprises the transaction data whereas block header contains a randomly generated 32-bit whole numbers called nonce and a 256-bit hash value that points to the previous block (parent block) as shown in Fig. 1. The first block in the chain which has no parent block is called the genesis block.

Block header also includes:

(i)Block version: indicates which set of block validation rules to follow.

(ii)Merkle tree root hash: Transactions within a block are not necessarily stored on first come first serve basis. It is organized in a merkle tree structure. It contains the hash value of all the transactions in the block.

(iii) Timestamp: current time as seconds in universal time since January 1, 1970.

(iv) nBits: target threshold of a valid block hash.

After the digital signatures of blocks are verified and all the transactions are validated, the block is then published and appended to the chain.

B. Identities:

An identity consists of a pair of public-private keys [9]. Most implementations of the blockchain, particularly public blockchain like bitcoin and ethereum use public key cryptography to maintain a certain level of anonymity. Both keys can be used to encrypt data that you generate. Data encrypted with public key can only be decrypted with private key and vice versa. Private key are used as a digital signature whereas public key is considered as the address. The typical digital signature is involved with two phases: signing phase and verification phase. Suppose, if someone wants to send data over the network then in signing phase data is encrypted with their private key and the signed data can only be decrypted by public key. Now in verification phase, the receiving node verify the sending node's public key. You can create as many identities you want and can store them in a digital wallet. But this also creates a challenge of malicious nodes, especially in public blockchains.

C. Nodes:

One of the most important aspects of Blockchain technology is decentralization. Nodes can be considered as an electronic device spread across the network that maintains copies of blockchain. Nodes are responsible for generating and verifying the transaction. After validation of block is done, they broadcast the block over the network. The nodes that create the blocks are called miners and the process is called mining.

IV. CLASSIFICATION

Blockchain technology can be classified in a number of ways, ie, on the basis of data it can take or the level of accessibility and data transparency: Here we discuss the latter criteria in detail.

Blockchain categorized on the basis of level of accessibility and data transparency takes into account the facts that to whom the data is visible and to what degree it is. On this criterion, blockchain can be broadly classified into three categories:

A. *Public Blockchain*:

In this classification, participation of nodes is not restricted and anyone can join the network. In this, a node is responsible for providing intermediation services like transaction verification to the client users. Since anyone can participate and all the participants are pseudonymous, it leads to a large number of nodes and makes the system prone to inconsistent data. Hence a complex consensus mechanism is required to maintain a unified chain records kept by all the nodes. It also increases the time required in transaction processing. In a public blockchain, nodes are called miners and all of them are given equal opportunity to read or write blocks. Most cryptocurrencies rely on public blockchain for their mechanism. E.g., Bitcoin, Litecoin, etc.

B. *Private Blockchain*:

In private blockchains, access to blockchain is restricted to a few trusted people only. This is mostly used for enterprises and allows them to implement distributed ledger systems without making data public. In private blockchains all the validation decisions are made by a single authority only. To determine the trustworthy validator, vetting is performed. No other complex consensus mechanisms are required since the final data is shared or broadcast by one node only.

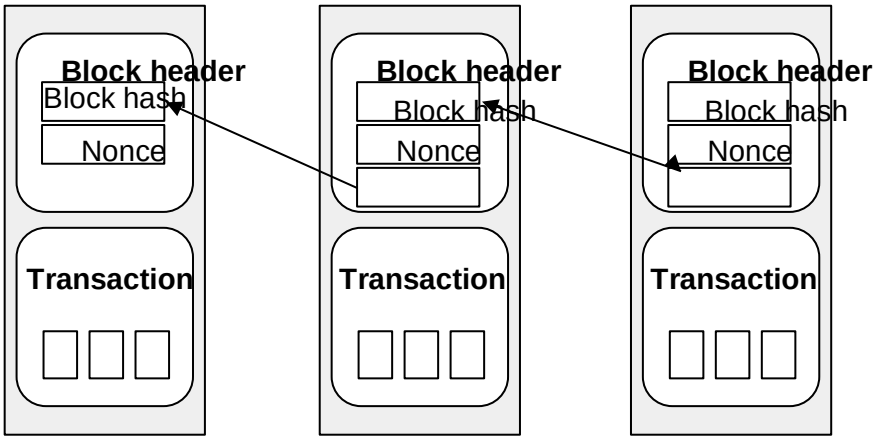


Fig. 1 Private Blockhain

C. Permissioned Blockchain:

This classification can be considered to be a combination of public as well as private blockchain. In this, like a public blockchain, data is visible to everyone but, like a private blockchain, access is controlled by just one authority. The consensus process is pretty simple because just like a private blockchain, permission is required for a node to participate in one such blockchain. Some enterprise focussed cryptocurrencies use permissioned blockchains like Ripple, Corda, Hyperledger fabric [6], etc.

V. WORKING OF BLOCKCHAIN

It is the unique and prominent working structure of the blockchain technology that contributes to its massive success and strong character. The basic building parts of a blockchain are blocks.

A block is basically an electronic entity which contains bits of data, nature of which depends upon the application blockchain is being used in. But there are some fundamental pieces of data in blockchain, that characterize a block in a network.

The first data in a block is the data associated with the properties, the blockchain is being used for. For example, in the application of blockchain in the financial sector, this data is the details of the transaction the block represents. This includes the initiator and receiver of the transaction, date/ timestamp of the transaction as well as some other details.

The other key aspect data that exists in a blockchain is the hash code of the previous block in the chain. A hashcode is a code, generated using a hash function. A hash function is an algorithmic/ mathematical function, that takes in a variable length input and gives out a constant length output. This hash function is the building block of the security concept of blockchain technology. Standard hash functions provide immutability in the fact that even a slight change in the input string, changing the case of a single letter for instance, drastically changes the output string of the function. Thus, when a block consists of the hash code of the block preceding this one in the chain, it results in the formation of a secure chain. For example, if a person tries to tamper with a block's data, he/ she will have to change the hash code of the previous block to, which in turn requires a change in the hash code of the previous block, and so on. This is why blockchain technology is one of the most trusted technology across various sectors.

When one such blockchain is developed for a single purpose, it is usually distributed to multiple people across the internet. Each recipient of this blockchain is called a node, and every node contains their own copy of the same blockchain. This distributed aspect of the blockchain provides it the characteristics of decentralization and transparency.

When the blockchain is shared to multiple users across the internet, the concept of one single hosting spot of data, centralization is eradicated, and rather, everyone contains their own copy of the blockchain. Thus, there is no chance of a single point of failure, or that of the client-server mechanism. Thus, there is no single authority providing details to the participants across the networks, or accepting input from them. Every individual or node present on the network can view or add blocks to the network.

Any blockchain, in general, consists of mainly 3 components.

1. User Applications - This component includes the wallet or any service in that matter, which generates or initiates a transaction.
2. Blockchain Network - A Peer-to-Peer (P2P) Network, is the part of blockchain where all the transaction processing is done using any of the suitable consensus. Blockchain Network, Data and Transaction Processing, as discussed below in detail, are what forms the basis of the blockchain network.
3. Database - This stores the data regarding every aspect of the transaction in form of 256 binary bits or 64 hexadecimal bits.

A. Blockchain Network and Data Processing:

Data transmission is simple in centralized blockchains. If you're using a relatively centralized blockchain with a limited number of nodes like Ripple, this process is also easy. You just broadcast the data to all the nodes simultaneously and they'll receive and process the data in a relatively short time frame. By contrast in a distributed system data usually flows via a peer to peer network. In a distributed system, unlike a centralized system there are thousands of nodes so instead of sending it to the entire network, miners broadcast the transaction data to a few miners that are closest to them in terms of network latency. The miners will propagate the data on the network using what's called a gossip protocol. It is a more efficient way for transaction propagation.

B. Transaction Processing:

Many blocks use a scripting language to process incoming data. Any client application can generate a transaction script, i.e., a programming code for the blockchain node to execute. It's a simple script with some inputs and outputs. The input has a hash pointer containing the hash of a previous transaction whereas output consists of the hash of the next as shown in the Fig.2. Client publishes the transaction throughout the network of nodes. Nodes then execute the script to verify the validity of the transaction.

VI. CONSENSUS MODEL

In blockchain, Consensus protocols exist to address the Byzantine general's problem. Once a system becomes decentralized, a major downside of this setup is that different nodes are likely to receive different information. Since transactions are not propagated to all nodes at the same time due to network latency, it is possible that at any point of time, different nodes will have different data records.

Consensus is an algorithmic [12] process to ensure that there exists a single copy of records shared by all nodes but it doesn't guarantee that data copy is more secure than its centralized counterpart. The goal is to reduce these different data copies into one unified copy shared by all the nodes. This can easily be achieved by simply imitating a centralized system.

A. CENTRALIZED CONSENSUS MODEL:

The idea behind this consensus model is that we designate a central node and whatever data the central node has is going to be the official copy. In fixed intervals, central node collects all the transactions it has received into blocks and then validates them. The final copy is then broadcast to all the nodes. Once it's posted, the rest of the nodes simply assume a much more passive role by verifying the crypto signatures for the transactions in the block and accept them. A simple variation of this approach is the majority vote. In this process, the nodes simply post their transactions and transactions that have been heard by a majority of the nodes are automatically accepted onto the blockchain. This central node setting is unbeatable in terms of efficiency and since only central node takes the validation decision, it is fast as well. There's no need for any mining or complex computation. This kind of consensus model is often used for private and enterprise focussed blockchain for e.g., Ripple.

But this kind of defeats the purpose of blockchain as a distributed system as in this case all power is in the hands of central node only. Only the Central node can make decisions about validation of a block and other nodes act more passive. Also vetting is required to decide a trustworthy validator. This vetting process has to take place off the blockchain using traditional legal means which is quite expensive and consequently, the central validator process is not applicable to public blockchains where the nodes are more or less equal and all of them are anonymous.

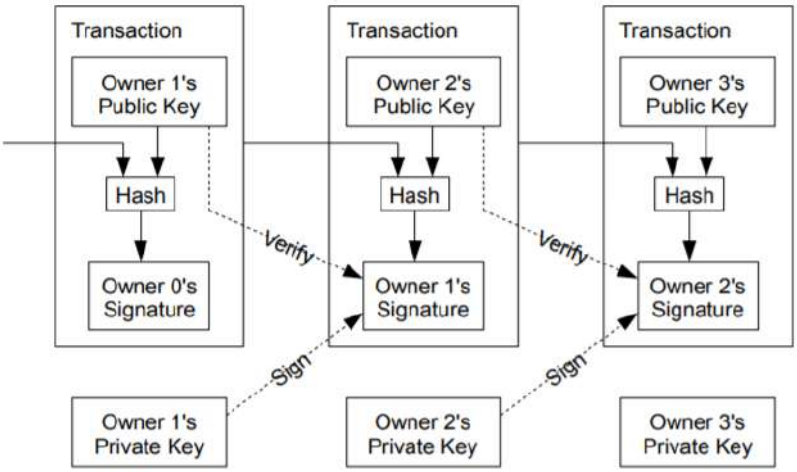


Fig.2 Centralized Consensus Model

B. RANDOM CONSENSUS MODEL:

As we see in the central consensus model, all the power of validation of a block lies in the hands of central validator at every interval and this creates the need of vetting. In random consensus model we consider what if vetting is impossible, like in public blockchains where all the nodes are equally anonymous to each other. So, in this case, we just randomly select a node and let it broadcasts its pending block. If all the transactions in the broadcast are valid then other nodes will accept the block and add to the chain. But if it's not valid then another node is randomly selected and the block is rebuilt. So, at the end of every interval, we're going to throw a dart at the board of nodes and randomly select one node to post this block. And here again the rest of the nodes assume a more passive role and simply verify the validity of transactions in the block, like having the right crypto signatures etc. But this mechanism only ensures that every node in the network would have the same chain of data blocks. It doesn't address the issue of malicious nodes. If a fake node gets picked in a period then it can compromise the system. Since in public blockchains all nodes are pseudonymous and one can create as many fake identities, it is easy to carry out double spend attacks. Hence to create a more robust consensus mechanism, in the process of randomly selecting the node, we can weigh it by something that is hard to fake,i.e., computing power or total wealth.

VII. CONSENSUS ALGORITHMS

There are various consensus algorithms developed out there and all have their own pros and cons. Here, we have discussed a few of them in detail.

A. Proof of Work:

PoW (Proof-of-work) or Mining is a consensus mechanism used in the Bitcoin network [1]. In this strategy, nodes are weighed based on their computing power. Ergo, this makes a successful consensus attack economically infeasible and makes the system less vulnerable to Sybil attack. The proof-of-work consensus is by far the most widely used consensus protocol.
The most widely used proof-of-work consensus is based on SHA-256 and was introduced as a part of Bitcoin. Others include Scrypt, SHA-3, scrypt-jane, scrypt-n, etc.

The main idea behind this algorithm is that nodes have to compete to solve complex mathematical problems in order to mine the next block. This mathematical puzzle requires a lot of computational power. The solution to the problem is called hash value of the block header. The nodes that calculate the hash values are called miners. Miners append a random nonce value to the block to find a fixed hash of leading 0's. The nonce is 32 bits and hash is 256 bits. This gives roughly four billion possible nonce-hash combinations that must be mined before the right one is found. Brute forcing is only an option in this case. When a miner finally calculates the required hash value, it broadcasts the block to the whole network along with nonce and other nodes verify the solution and block validity. After validation, the block is appended to the blockchain.

But this approach has its own limitations. If a controlling entity owns 51% or more than 51% of nodes in the network, the entity can corrupt the blockchain by gaining the majority of the network. Also, since brute forcing is the only option, it is a time-consuming process as well. Miners consume high amounts of computing power in order to find the solution to the hard-mathematical puzzle. It leads to a waste of precious resources (money, energy, space, hardware).

B. Proof of Stake:

Proof of stake is considered to be the most common alternative to POW algorithm. Many blockchains adopt PoW [2] at the beginning and transform to PoS gradually. For instance, ethereum has shifted from POW(Ethash) to POS(Casper).

Proof of stake(POS) or virtual mining weighs the node on basis of the amount of cryptocurrency you invest. It is the more direct form of mining. In this type of consensus algorithm, instead of investing in expensive hardware to solve a complex puzzle, miners stake their identities directly with cryptocurrencies. Consequently, miners become direct stakeholders of cryptocurrency. This selection algorithm combines the quantity of stake (amount of cryptocurrency) with other factors (like coin-age based selection, randomization process) to make the selection fair to everyone on the network. In PoS algorithm Nodes make transactions. The PoS algorithm puts all these transactions in a pool. Each node places a cryptocurrency stake to bet on its block. Probability of a node being selected to post the block is directly proportional to the miners bet. The selected node then broadcasts the block across the network and other nodes verify the blocks validity. Once the validation is done the block is added to the chain. The miner that forges the block is rewarded with a transaction fee. This approach does not require miners to solve any complex mathematical problems, thus consuming less time. This also increases the efficiency as well as significantly lessens the energy and resource requirements of the system.

The major drawback of this approach is that If a group of validator candidates combine and own a significant share of total cryptocurrency, they will have more chances of becoming validators. Increased chances lead to increased selections, which lead to more and more forging reward earning, which lead to owning a huge currency share. Therefore, in the long run, the system can become significantly centralized thereby defeating the purpose of decentralization in blockchain.

B. Practical Byzantine Fault Tolerance (pBFT):

pBFT consensus algorithm is designed to work in an asynchronous system. The Byzantine Fault Tolerance (BFT) [10] defines the dependability of distributed computing [13] [14] systems where components may fail and result in imperfect information. Achieving BFT is one of the most difficult challenges addressed by blockchain technology. The objective of a BFT mechanism is to safeguard against the system failures by employing collective decision making(both – correct and faulty nodes) which aims to reduce the influence of the faulty nodes.

pBFT [15] tries to provide a practical Byzantine state machine replication that can work even when malicious nodes are operating in the system. Nodes in a pBFT enabled distributed system are sequentially ordered with one node being the primary(or the leader node) and others referred to as secondary(or the backup nodes). The goal is that all honest nodes help in reaching a consensus regarding the state of the system using the majority rule. If needed, a majority of the honest nodes can vote on the legitimacy of the current leading node and replace it with the next leading node in line.

A practical Byzantine Fault Tolerant system can function on the condition that the maximum number of malicious nodes must not be greater than or equal to one-third of all the nodes in the system.

pBFT consensus rounds are broken into 4 phases: The client sends a request to the primary node which broadcasts the request to all the secondary nodes. The nodes perform the service requested and then send back a reply to the client.

The request is served successfully when the client receives ‘m+1’ replies from different nodes in the network with the same result, where m is the maximum number of faulty nodes allowed.

pBFT is an energy efficient approach to achieve consensus as it does not require any complex calculations. Every node takes part in responding to the request by client and hence every node can be incentivized leading to low variance in rewarding the nodes that help in decision making.

The pBFT consensus model works efficiently only when the number of nodes in the distributed network is small due to the high communication overhead that increases exponentially with every extra node in the network. The pBFT mechanisms are susceptible to Sybil attacks, where one entity(party) controls many identities.

C. Delegated Proof of Stake:

Delegated Proof of Stake (DPoS) is a consensus algorithm developed to secure a blockchain by ensuring representation of transactions within it. DPoS is designed as implementation of technology-based democracy to protect the system from malicious nodes. Stakeholders elect their delegates to generate and validate blocks. With significantly fewer nodes to validate the block, the block could be confirmed quickly, leading to the quick confirmation of transactions.

Meanwhile, the parameters of the network such as block size and block intervals could be tuned by delegates. Additionally, users need not to worry about the dishonest delegates as they could be voted out easily. DPoS is largely considered to be the most decentralized approach to consensus mechanism. It also is an energy efficient and time saving method that has strong protection from double spend attacks. Cryptocurrencies that use DPoS include Lisk, Steem, EOS and BitShares.

VIII. CHALLENGES OF BLOCKCHAIN TECHNOLOGY

In Spite of having such a remarkable performance over the years, blockchain technology still faces some major challenges in the industry. A few of those challenges are implementation dependent and some are for the general working of the technology. Both of these types of challenges are collectively discussed below:

A. Scalability:

Blockchain technology uses the concepts of decentralization and transparency as its core features. Thus, it is obvious that it supports a large network of nodes. This large number of simultaneous users, in turn, increase the cost and reduce the throughput of the blockchain [5]. They make the blockchain bulky and as a result decreases the number of transactions.

These scalability issues can be addressed via majorly two modes:

1. Lighting Network - It is a two-layer protocol that offers an off-chain settlement among the participants which aims at processing those transactions first which have low cost and faster processing time.
2. Sharding - It is a method in which subsets of nodes are grouped into smaller networks, often referred to as shards. Shards are responsible for a transaction specific to a particular shard.

B. *Security:*

Blockchain as a whole is a very secure and immutable record system [11]. Also, on the other hand, if a few users are involved in false or malpractices, this could put the whole blockchain in jeopardy. Thus, this does not guarantee transactional safety as the transaction details and balances publicly visible.

C. *Selfish Intents:*

Some nodes on the network, with selfish intentions, do not show their mined blocks to the network. This private branch is only made public when certain requirements are fulfilled. This leads to the honest miners waste their time on mining the already mined blocks, with the selfish miners [8] having access to majority or all parts of the blockchain.

This issue can be solved by an approach where all the users are allowed to follow a branch. Using random beacons and time stamps, honest miners can select fresh blocks. This poses another issue, the fact that timestamps can be forged, which can be overcome by issuing a specific time for a block to be generated and accepted in the network.

IX. APPLICATIONS OF BLOCKCHAIN TECHNOLOGY

Blockchain has received immense popularity and support over the years it has been around and thus there have been numerous applications concerning various sectors in the industry. Following are discussed some of the applications of blockchain technology in various parts of the world.

1. Food Sector:

The food sector has seen some tremendous approaches with blockchain as the market suffers from its own problems of traceability and sustainability. The fact that blockchains are immutable and distributed ledger systems, gives the food producing and quality assurance agencies an easy hold on checking the source of bad food products and the consumers to track the source and path of their commodity, in order to ensure its sustainability.

Small level farmers face the problem of getting capital for their produce and the data sharing and distributed properties of blockchain help them in obtaining funds from trusted investors.

2. Election Commission:

India, along with many other countries, has invested in the blockchain technology to smoothen out and improve the election procedures. The fact that every block ever created in a blockchain is unique, with its own unique hash code, and is thus immutable, that is, it cannot be tampered with, provides the level of trust and security that an activity as important and prime as election requires. Since everyone on the network of a block chain can view all the contents of each and every block, this provides the transparency factor that is a key factor in the election procedure.

3. *Crypto currency:*

Bitcoin has been recognized as one of the world's most popular crypto currency and it has got blockchain to owe to it's massive success. The three pillars of blockchain technology: Immutability, Decentralization and Transparency gives Bitcoin and other cryptocurrencies all the desired features that are responsible for it's huge outreach.

4. *Health sector:*

Storing patients records unhindered by outside influence yet available for all concerned people to easily keep track of everything makes blockchain a boon for the health industry.

5. *Smart contracts and Notary:*
The fact that once a block has been registered in any given blockchain, it or it's data cannot be tampered with just makes blockchain a suitable technology behind smart contracts [4] [7] and e-notary.

X. FUTURE ASPECTS OF BLOCKCHAIN TECHNOLOGY

Blockchain, as discussed in this paper, has a lot laid out for its future days, and we are sure to see it getting a lot of advances in the coming days and years. Following are discussed some of the future trends that can be included with the blockchain technology to give it the further push it requires.

A. Big Data Analytics:

Blockchain can be combined with Big Data as its model is that of the decentralized and distributed along with being secure. Thus, it can ensure that the data stored in it is original. A study of the transactions taking place on the blockchain can provide the data analytics with the details and information required.

B. Centralization.

One of the most major technology patterns is the centralization concept of the data, giving authority over the data to a single entity/server. Blockchain technology, in its decentralization approach, stands tall to give this conventional concept of the internet a hard challenge. Despite this feature miners are centralised in the mining pool. According to a recent report, the top 5 mining pools together own larger than 51% of the hash value in a bitcoin network. But blockchain is not intended to serve only a part of the organisation. Hence, this is a problem blockchain has to overcome in order to go against the centralization concept.

XI. CONCLUSION

Blockchain is a revolutionary technology, in its novice state. It has certain problems but not without a lot of potential. It has the capability to transform the industry with its leading aspects of audibility, decentralization, persistence and anonymity. This paper presents a basic overview of blockchain which comprises blockchain architecture, algorithms and characteristics of blockchain. Furthermore, its applications and challenges are also included here. Blockchain applications are increasing tremendously and it has a huge potential to come out as technology of its time.

REFERENCES

[1] Satoshi Nakamoto ,"*Bitcoin: A Peer-to-Peer Electronic Cash System*"[2] Zibin Zheng, Shaoan Xie, Hongning Dai, Xiangping Chen, and Huaimin Wang ,"*An Overview of Blockchain Technology:Architecture, Consensus, and Future Trends*", in press

[2] Zibin Zheng, Shaoan Xie, Hongning Dai, Xiangping Chen, and Huaimin Wang ,"An Overview of Blockchain Technology:Architecture,

Consensus, and Future Trends", in press

[3] NRI, "*Survey on blockchain technologies and related services*," Tech. Rep., 2015. [Online]. Available: http://www.meti.go.jp/english/press/ 2016/pdf/0531 01f.pdf

[4] V. Buterin, "*A next-generation smart contract and decentralized application platform*," white paper, 2014.

[5] M. Crosby, P. Pattanayak, S. Verma, V. Kalyanaraman, "*Blockchain technology: Beyond bitcoin*"

[6] "*Hyperledger project*," 2015. [Online]. Available: https://www. hyperledger.org/

[7] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "*Hawk: The blockchain model of cryptography and privacy-preserving smart contracts*," in Proceedings of IEEE Symposium on Security and Privacy (SP), San Jose, CA, USA, 2016, pp. 839–858.

[8] I. Eyal and E. G. Sirer, "*Majority is not enough: Bitcoin mining is vulnerable*," in Proceedings of International Conference on Financial Cryptography and Data Security, Berlin, Heidelberg, 2014, pp. 436– 454.

[9] A. Biryukov, D. Khovratovich, and I. Pustogarov, "*Deanonymisation of clients in bitcoin p2p network*," in Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, New York, NY, USA, 2014, pp. 15–29.

[10] L. Lamport, R. Shostak, and M. Pease, "*The byzantine generals problem*," ACM Transactions on Programming Languages and Systems (TOPLAS), vol. 4, no. 3, pp. 382–401, 1982.

[11] Dr. Divyakant Meva "*Issues and Challenges with Blockchain: A Survey*" in INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING, December 2018, in press

[12] D. Kraft, "*Difficulty control for blockchain-based consensus systems*," Peer-to-Peer Networking and Applications, vol. 9, no. 2, pp. 397–413, 2016.

[13] Sujaya Maiyya, Victor Zakhary, Divyakant Agrawal and Amr El Abbadi, "*Database and Distributed Computing Fundamentals for Scalable, Fault-tolerant, and Consistent Maintenance of Blockchains*" in Proceedings of the VLDB Endowment, Volume 11, Issue 1

[14] M. Castro, B. Liskov, et al. Practical byzantine fault tolerance. In OSDI, volume 99, pages 173–186, 1999.

[15] Xu Hao, Long Yu, Liu Zhiqiang, Liu Zhen and Gu Dawu, "*Dynamic Practical Byzantine Fault Tolerance*" in Proceedings of 2018 IEEE Conference on Communications and Network Security (CNS),

Floral Ailment Identification

Rushali Khanna[1], Aman Sharma[2], Prakash Choudhary[3], Gunjan Chugh[4]

[1,2,3] *Student B-Tech-IT, ADGITM, GGSIPU University, New Delhi, India*
[1]rushalikhanna30@gmail.com
[2]amans5517@gmail.com
[3]pcchoudhary77@gmail.com

[4]*Assistant Professor, IT Engineering, ADGITM, GGSIPU University, New Delhi, India*
[4]chugh.gunjan8917@gmail.com

***Abstract* – The agronomical sector constitutes a major sector in the society when compared to other occupations. About 60-70% of Indian population depends on the agricultural industry. This sector is ironically even prone to the maximum amount of diseases and thus incurred loss. Diseases could be a joint cause of weeds and arthropods whose percentage has increased about an alarming rate of 30%. Bacterial and fungal infections constitute the main reason for these diseases. Our paper presents a CNN model and an algorithm to detect such diseases through the leaf of the crop and thus provide the viable alternatives and solutions that are possible. Our model aims to read and understand all images and thus classify them either as diseased or healthy leaves. Furthermore, it also suggests the viable options one has when the plant has gotten effected by the disease. We are using InceptionV3 algorithm and an 'Adam' Optimiser.**

***Keywords* - Convolutional neural network, Deep Learning, InceptionV3, apple, potato, strawberry, corn, grape, Softmax, Adam, PlantVillage.**

I. INTRODUCTION

The first and foremost pedestal on which the society stands is its farmers and the agronomical sector. For many years India is one of the major countries involved in the agricultural sector [1]. This industry is an inevitable part of our society due to its stellar contributions. Among the various crops grown in India, potato, apple, grape strawberry, and corn are the plants we have worked upon in this paper. In India's diverse range of subtropical climate, growing these fruits and vegetables yields fruitful results. When a particular specimen is defected from its actual well-being due to foreign substance it is classified as a diseased plant. The disease usually strains the crop and disrupts the vitality of the plant and makes it less valuable. The diseases that a plant might be suffering from may also have a very big environmental cause that is the changing season. Pathogens are one of the main reasons due to which disease spread increases many folds. Pathogens are directly followed by incorrect variety of the crop grown in a particular season. One of the common diseases faced by potatoes are Early and Late Blight [1]. One of the diseases apples face is apple cedar rust [2]. Corn and strawberries face northern leaf blight and leaf scorch respectively [2]. Grapes tend to face black rot [2]. They can have adverse effects on plants and agricultural lands. These diseases can be spotted on the plant leaves but it's is very time consuming if spotting is done manually. Thus, newer reforms are required.

There are many ways to detect objects and its special features from their image with the help of image processing and computer vision. Convolutional Neural Nets still stand to be the most common way approach for deep learning [3]. The model we have come up with is a machine learning model that is trained to learn the disease of a plant and then tested to check its accuracy. Also, this model provides a fool-proof support to the user with a list of various viable options and solutions according to the disease it detected.

II. CONVOLUTION NEURAL NETWORKS

Convolutional neural networks are learning techniques applied to a set of images which can be in a varied group of formats. This is done to understand and realise the images [3]. They are also known as shift invariant. Their weight sharing architecture and translation invariance characteristics help the neural network models to study and analyse the images. Henceforth the coining of the term - space invariant neural networks. Recommender system, image and video recognition are also some viable applications of CNN. Even the medical image analysis, image classification, natural language processing(NLP), brain-computer interfaces that

communicate between them, and financial time series are some of the applications of CNN. Varied fields still possess the feature to be exploited by neural networks that today stand absolutely undiscovered.

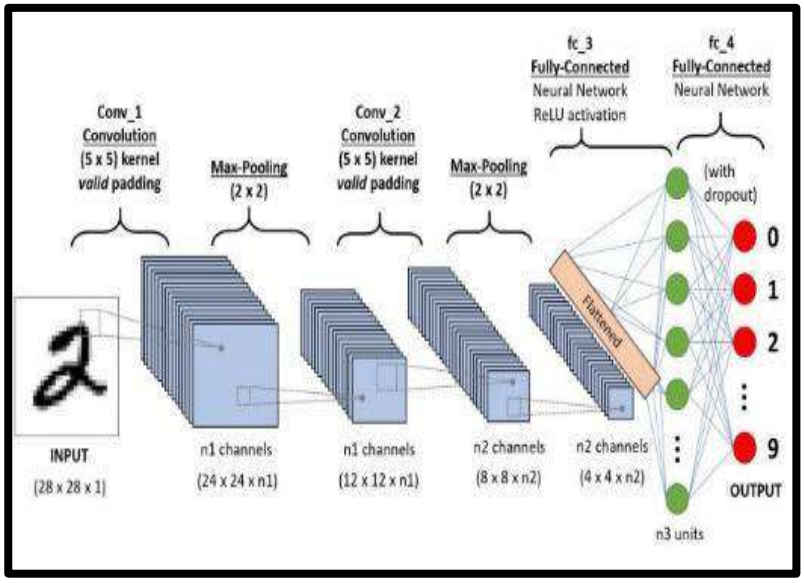


Fig 1. Convolution Neural Network [4]

III. RELATED WORKS

Detecting a disease for a farmer itself is a very big task and doing so manually makes it more tedious. There are many engineers already leading the path towards disease free plants and encouraging more and more farmer to follow the foot-steps. Detecting diseases based research is as follows: D. Tiwari et. al. [1] used transfer learning and various pre-trained model on the potato leaves image dataset and concluded that VGG 19 gave the optimal accuracy of 97.8% compared to 92% accuracy of backpropagation neural network and 95% accuracy of support vector machine. Melike Sardogan et. al. [5] in their research has tried to use Convolutional Neural Network with Learning Vector Quantization algorithm to detect and classification of tomato leaf diseases.three input matrices are obtained from 500 images of tomato plants for R,G and B channel. This is done for every image in the dataset at the start of the process. Disease detection and image classification has been performed by LVQ algorithm. In the process of classifying the tomato leaf diseases, clubbing CNN with the profound LVQ algorithm yielded accurate and least errored results. Halil Durmus et.al. [6] utilised the famous architectures in their paper namely- AlexNet and SqueezeNet on the hardware Nvidia Jetson Tx1 to perform their research. However, he also came up with a conclusion that AlexNet develops an extremely bulky model which isn't compatible on mobile phones. Its size is 227.6 Mbyte whereas that of SqueezeNet is 2.9 Mbytes. Thus, the model built by SqueezeNet is easily compatible on cellular devices. Also, it doesn't provide any improvement in the inference time. Henceforth SqueezeNet yields most accurate results when working on the Nvidia Jetson Tx1.U.Atila et.al. [7] Architecture for cellular devices. They tried to compare state-of-the-art deep neural net architectures with the EfficientNet deep learning architecture to detect plant leaf disease detection in the Google Cloud Environment. Other Cnn architectures tend to give a lower accuracy however EfficientNet acrchitecture blesses us with 99.97% accuracy and almost a negligible error rate. The precision of CNN offered is quite less as compared to EfficientNet.

IV. DATASET DESCRIPTION

A deep learning model has a basic building block without which it cannot be built and would collapse. This is the dataset that governs the main entity of the neural net. The chosen given dataset is used to fulfil two main objectives which are training of the neural network model and testing of the neural network model on the basis of the training. The dataset that this model has employed is the Plant Village Dataset [2] that is taken from Kaggle. A safe and secure platform for all machine learning enthusiasts is Kaggle. It enables on to access millions of tons of data which is compiled in an organised manner so one can use it for machine and deep

learning ventures [1]. In this dataset there are about 20000 images of leaves of potato, apple, strawberry, corn, grape plants and many more. The dataset contains jpg/png format of images. This dataset contains both kind of images, the healthy and the diseased leaves. Four-fifth of this dataset goes into training for healthy leaves and the one-fifth of the set goes into the testing procedure; and so is the case for each of the diseased leaves. Henceforth, the model created will be trained on a larger dataset and tested on a comparatively lesser dataset to achieve the highest accuracy during prediction and detection. The dataset has been properly organised into files and directories. Each plant has a separate folder assigned to it which holds all the images of that plant. This folder has 2 directories that help to distinguish the diseased leaves from the healthy ones.

The subset of the dataset we are using is the Potato plant, apple plant, strawberry plant, corn plant and grape plant dataset. In this dataset there are early blight leaf images that account to 1000 images, late blight leaf images that account to 1000 and 152 images of heathy leaves for potato. 1000 images for leaf scotch and 450 healthy leaves for strawberry. 1000 images for black rot and 400 healthy leaves for grapes. 250 images for cedar rust and 250 healthy leaves for apple. 950 images of northern leaf blight and 700 healthy leaves for corn. Some images from the dataset are shown below:

Potato Leaf Apple leaf Strawberry leaf Corn leaf Grape leaf

Fig 2. Snippet of samples used [2]

V. PLATFORM UTILIZED

The hardware used in the paper is a laptop with the specification of 8 GB DDR4-2133 SDRAM (1 x 8 GB) with Intel Core i5-7200U and 1 TB hard drive with 5400 rpm. Here Matplotlib library is used to plot graphs related to the model's performance as well as data visualization. Matplotlib is an open-source python library used as a graphing and plotting tool. For Machine learning we have used TensorFlow and Keras library of python. TensorFlow and Keras are some of the open-source software library for deep neural nets [8]

VI. PROPOSED APPROACH

A convolutional based neural network in our project is being deployed to train a pre-processed dataset. The model received is then examined on a completely new lesser piece of dataset to test the accuracy of the model. The model is also capable of providing the user with viable alternatives and optimal situation solutions [9], [10], [11].

A. Feature Extraction

Inception V3 architecture [8] is the basic foundation step of the neural network model. This model enables us to presume feature extraction to the dataset. Google and their team of distinguished members have developed the Inception V3 algorithm. Convolutions, dropouts, max pooling, concatenations, average pooling and fully connected layers are the basic retributions that help the model stand up and function. Activations inputs are applied with batch norm throughout the layers contained in the model [8]. This process helps the images to distinguish between necessary and irrelevant data pixels present in the image and thus learn only the most minute and integral of the details.

This technique referred to as feature extraction adds an advantage to the programmer's bag by highlighting the important characteristics of the image and keeping them intact for future references and interpretation [3].

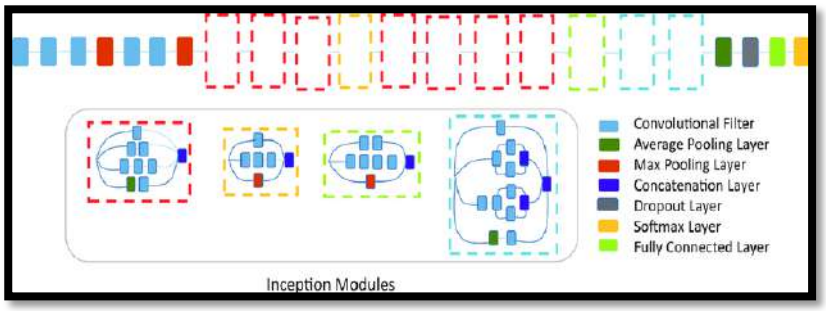


Fig 3. Inception V3 Architecture [12]

B. Classification

Classification process has been performed by the convolutional based neural nets. The feature extraction done by the Inception V3 [8] beforehand comes in handy here. It helps us to train the model to efficiently detect presence of a disease in the leaf images. However, classification isn't a fool-proof method thus, to reduce training time and quickly converge the loss, Adam optimiser is put to utility. Adam optimiser is an extension of the Stochastic Gradient Descent that is more commonly used for computer vision nowadays [13]. There is also a SoftMax cross entropy function that comes into play when classification has to be performed in different labels. The softmax activation function is a function that converts a vector of n real values (whose sum is not 1) into a vector of n real values that totals up to 1 [14]. The softmax transforms them into values even if the input it gets is in the form of numbers. These values resulting from transformation may be between 0 and 1. These are helpful in interpreting them as probabilities.

C. Model

The convolutional neural network model contains an input layer where all images are sent as an input in the start. After which the images are fed to the inception v3 architecture [8] which performs the required feature extraction. These feature extracted images are the forwarded to deep neural nets which contain many hidden layers that classify our images on pre-trained knowledge. Finally, the read data is then passed over to the ultimate layer that analyses the understanding of the model till the last stage. The model has a set of layers that are first initialised and after that they are compiled together using functions from the TensorFlow library [8].

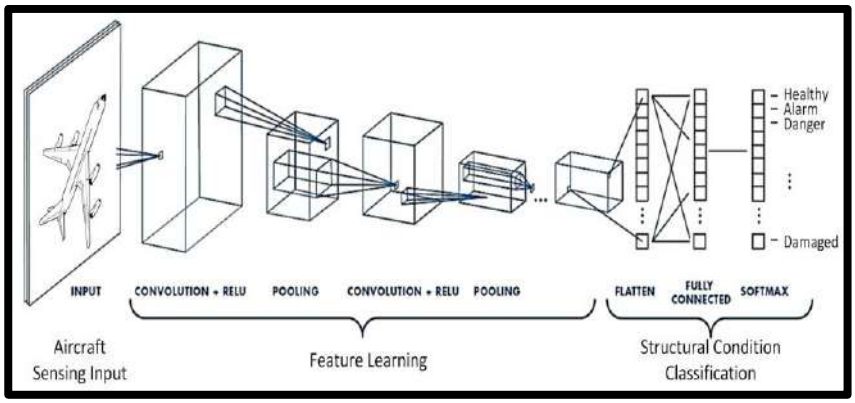


Fig 4. CNN Model [15]

VII. RESULTS

Plant Village dataset has been extracted from a free online platform Kaggle [2] which has about 1000 leaf images of early blight and 152 images of healthy plants. 1000 images for leaf scotch and 450 healthy leaves for strawberry. 1000 images for black rot

and 400 healthy leaves for grapes. 250 images for cedar rust and 250 healthy leaves for apple. 950 images of northern leaf blight and 700 healthy leaves for corn. As per the requirements of the model, the dataset is split up into 2 segments namely the training set and the test set. Training set comprises of 80% of the dataset while the test set comprises of 20% of the dataset. Feature extraction for this dataset has been done by a pre-trained model that is Inception V3[8]. Our model achieved an accuracy of 96% on the training and test set. Inception V3 as stated above yields the best and most accurate results with least error %

Other models present in the market have also been compared by us. It was found that Inception V3 architecture produced the best results. VGG16 algorithm produces an accuracy of 56% and RESNT50 produces an accuracy 92% [16]. The following accuracies had been developed by Utsav Vanodiya in his project namely – "Exploratory Data Analysis" using PlantVillageDataset.

Fig 5. VGG16 and RESNET accuracy [16]

The confusion matrix constructed to analyse the performance of the model has been presented below. It indicates the number of correct and incorrect predictions both sub divided into more categories such as true positive, false positive, true negative and false negative.

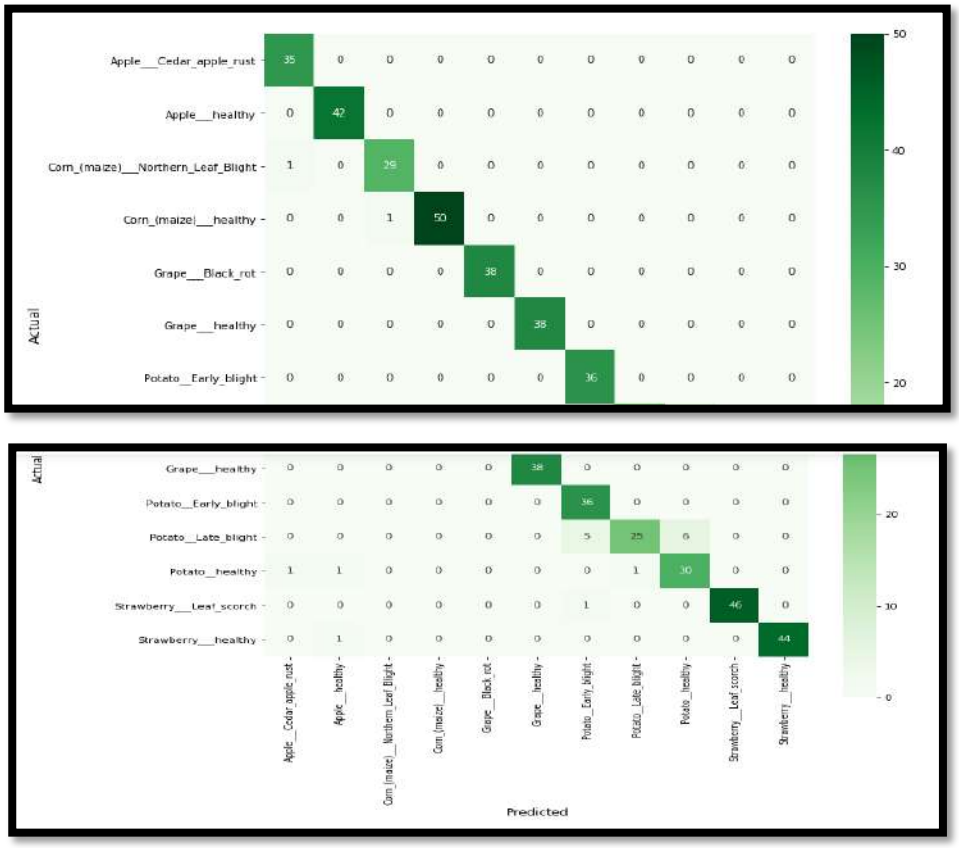


Fig 6. Model confusion matrix

Also the accuracy and the loss has been calculated using various graph representations. These graphs help to visualise the accuracy of the model and understand it in a better format than written numerical. Thus we have created graphs using matplotlib pyplot library to hep visualise our model.

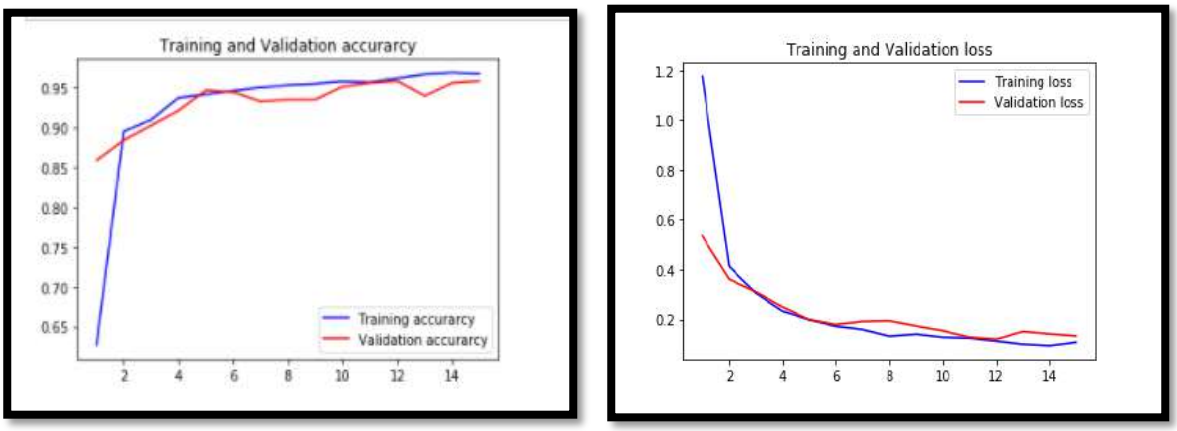


Fig 7. Model accuracy graph

VIII. CONCLUSION

In this research, we have made a CNN model with the help of Inception V3 architecture and Adam Optimiser to diagnose and classify disease of plants where we achieved an accuracy of 96% over the test dataset in classification. With the help of this CNN

model, an agricultural worker can build a system from which the plants health can be monitored and issues in the plants more efficiently, enhancing the crop yield and will detect and diagnose diseases in the early stages itself. This model not only detects the diseases but also suggests how to deal with the disease and what viable options or solutions the farmer has to correct the loss that has happened [9, 10, 11].

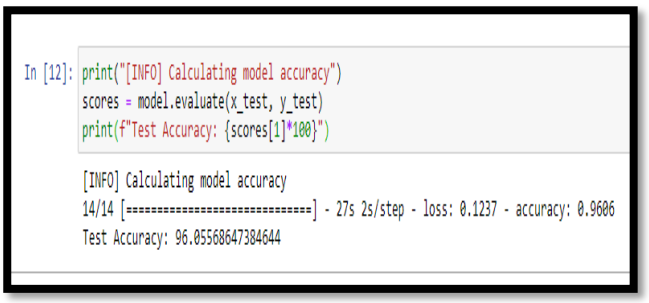


Fig 8. Model accuracy

IX. FUTURE SCOPE

In today's date, newer and better ways have been regularised to tackle diseases at the very minimal stage of its start. This helps to cause much less adverse effects to the plants and save them at the best conditions. However, if there is a delay in the process of detection of diseases it leads to a havoc situation in the plight of the plant as it leads to very poor produce and thus let the farmer to chemical alternative for manure. Only a detection of disease is not going to help the farmer from preventing all sorts of losses. Thus we propose to install this CNN model to all small and large scale farming lands to provide safe and secure growth to the crops and also timely coming up of solutions for the crops.

ACKNOWLEDGEMENT

We would like to thank our college Dr. Akhilesh Das Gupta Institute of Technology and Management for providing us with this opportunity to write a research paper along with our major project. We would also like to mention the continual effort and guidance provided to us by Ms. Gunjan Chugh. Without their contiguous support and help throughout the journey of the paper, it would never have been possible to accomplish such a heavy milestone.

REFERENCES

[1] D. Tiwari, Adem Tuncer, S. Bhardwaj M. Ashish, N. Gangwar, A. Sharma, S. Patel, "*Potato Leaf Diseases Detection Using Deep Learning*", International Conference on Intelligent Computing and Control Systems, 2020.
[2] Tairu Oluwafemi Emmanuel, " PlantVillage Dataset ", [online] Available: https://www.kaggle.com/emmarex/plantdisease
[3] Chee Seng Chan, Sue Han Lee, Simon Joseph Mayo, Paolo Remagnino, "*How Deep Learning Extracts and Learns Leaf Features for Plant Classification*", Pattern Recognition, 2017.
[4] Tabian, I.; Fu, H.; Sharif Khodaei, Z. A Convolutional Neural Network for Impact Detection and Characterization of Complex Composite Structures. Sensors 2019, 19, 4933.
[5] Adem Tuncer Melike Sardogan, and Yunus Ozen, "*Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm*", International Conference on Computer Science and Engineering, 2018.
[6] H. Durmus¸, E. O. Gunes ¨ ¸ and M. Kırcı, "*Disease detecti``on on the leaves of the tomato plants by using deep learning*", In Agro-Geoinformatics, IEEE 6th International Conference on, pp. 1-5, 2017.
[7] Ü. Atila, M. Uçar, K. Akyol, et al.,"*Plant leaf disease classification using EfficientNet deep learning model*", Ecological Informatics, 2019.
[8] Xiaoling Xia, Cui Xu and Bing Nan, "*Inception-v3 for flower classification*", 2017 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, 2017, pp. 783-787.
[9] Waterworth, K. (2020). What Is Black Rot On Grapes: Learn About Black Rot Grape Treatment [Online]. Available:https://www.gardeningknowhow.com/edible/fruits/grapes/black-rot-grape-treatment.htm
[10] Grant, A. (2020). Potato Early Blight Treatment – Managing Potatoes With Early Blight [Online]. Available: https://www.gardeningknowhow.com/edible/vegetables/potato/potato-early-blight-treatment.htm
[11] George, H. (2019). How to identify, prevent, and control cedar apple rust [Online]. Available: https://gardenerspath.com/how-to/disease-and-pests/cedar-apple-rust-control

[12] D. Yiming , S. Jae , K. Michael , T. Hari, H. Roy, J. Nathaniel, L. Dmytro, C. Timothy, A. Mariam, A. Carina, B. Spencer, F. Robert, H. Shih-ying, Z. Kelly, N. Lorenzo, S. Youngho, H. Randall, H. Miguel, H. Dexter, F. Benjamin. (2018). "*A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18 F-FDG PET of the Brain* Radiology." 290. 180958. 10.1148/radiol.2018180958.

[13] Kingma, Diederik & Ba, Jimmy. Adam: "*A Method for Stochastic Optimization. International Conference on Learning Representations*" , 2014 .

[14] Nwankpa C, Ijomah W, Gachagan A, Marshall S "*Activation functions: comparison of trends in practice and research for deep learning.*" Preprint arXiv:1811.03378, 2018.

[15] Saha, S. (2018). *A CNN sequence to classify handwritten digits* [Online]. Towards Data Science. Available: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

[16] Vanodiya, U. (2021). Exploratory Data Analysis [Online]. Available: https://www.kaggle.com/utsav7vanodiya/exploratory-data-analysis

Social Distancing Analyzer – Covid 19

Ikshit Bansal[1], Kartik Aggarwal[2], Girish[3], Naman Kumar[4], Gunjan Chugh[5]

[1,2,3,4] *Student B-Tech-IT, ADGITM, GGSIPU University, New Delhi, India*
[1] ikshitbansal1998@gmail.com
[2]kartikrocks110@gmail.com
[3]girish67678@gmail.com
[4]naman1999237@gmail.com

[5]*Assistant Professor, IT Dept., ADGITM, GGSIPU University, New Delhi, India*
[5]chugh.gunjan8917@gmail.com

***Abstract*— In the battle against COVID-19, social isolation has proved to be the most successful method of limiting the disease's spread. People are being advised to restrict their contact with one another in order to reduce the risk of the virus spreading through physical or near contact. In the past, AI/In-Depth Learning has shown promise in a variety of everyday problems. We will see a comprehensive explanation of how to track social distance in public places and workplaces using Python, Computer Vision, and deep learning in this proposed software. The social distance detection tool is used to ensure that social distancing policies are observed in public and at work. It can track people maintaining a safe distance from each other by analyzing real-time video streams from the camera. It can also be used to keep an eye on people at work, in warehouses, and in stores, and we can combine this tool with their security camera systems to see if people are keeping a safe distance from one another.**

***Keywords* – Social Distancing, Covid-19, Deep Learning, Machine Learning, Video Surveillance.**

I. INTRODUCTION

The novel generation of the coronavirus disease (COVID-19) was reported in end of December 2019 in Wuhan, China. After only a couple of months, the virus became a worldwide outbreak in 2020. On May 2020 the WHO announced things as pandemic [1,2]. The statistics by WHO on 8th October 2020 confirm 36 million infected people and a scary number of 1,056,000 deaths in 190 countries. With the growing trend of patients, there's still no effective cure or available treatment for the virus.

(a) Social distancing monitoring

(b) Accumulated infection risk (red-spots) due to multiple breaches of the social-distancing

Fig. 1 Social distancing, as shown in Figure 1a, refers to precautions to stop the spread of the disease

For several months, the planet Health Organisation believed that COVID-19 was only transmittable via droplets emitted when people sneeze or cough and therefore the virus doesn't linger within the air. However, on 8 July 2020, the WHO announced:
"**There is emerging evidence that COVID-19 is an airborne disease which will be spread by tiny particles remain in the air after people talk or breathe, especially in crowded and poorly ventilated settings" [2].**
Therefore, social distancing now claims to be even more important than thought before, and one among the simplest ways to prevent the spread of the disease additionally to wearing face masks. Almost all countries are now considering it as a compulsory practice. According to the defined requirements by the WHO, the minimum distance between individuals must be a minimum of 6 feet (1.8 m) so as to watch an adequate social distancing among the people [3]. Recent research has confirmed that folks with mild or no symptoms can also be carriers of the novel coronavirus infection. Therefore, it's important all individuals maintain controlled behaviours and observe social distancing. Many researches work like [4] have proved social-distancing as an efficient measure and a crucial inhibitor for limiting the spread of contagious diseases like H1N1, SARS, and COVID-19. Figure 2 depicts the effect of following social distancing guidelines to reduce the speed of covid infection transmission among individuals [5]. A wider normal curve with a shorter spike within the range of the health system service capacity makes it easier for patients to fight the virus by getting continuous and timely support from the health care organizations. Any unexpected sharp spike and rapid infection rate (such because the red curve in Figure 2), will cause service failure, and consequently, exponential growth within the number of fatalities.
During the COVID-19 pandemic, governments have tried to implement a spread of social distancing practices, like restricting travels, sealing borders, closing restaurants and bars, and alerting the people to maintain a distance of 1.6 to 2 m from one another. However, monitoring the quantity of infection spread and efficiency of the constraints isn't a simple task. People have to go out essential needs like food, medicines and other necessary tasks and jobs. Therefore, many other technology-based solutions and AI related research have tried to step in to assist the health and medical profession in coping with COVID-19 challenges and successful social distancing practices. These works differ from GPS-based patient localization and tracking to segmentation and crowd monitoring.

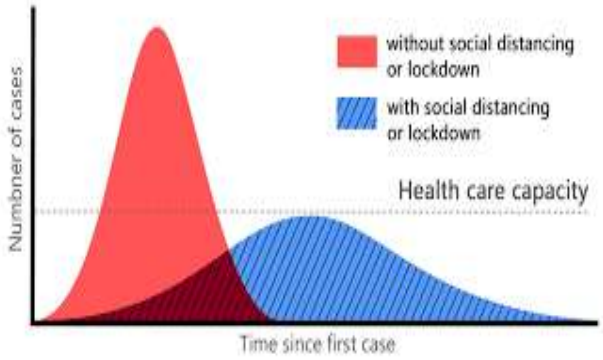


Fig. 2. Gaussian distribution of Covid-19 infection transmission rate for a given population, with and without social distancing norms.

II. TECHNOLOGY USED

This paper demonstrates the detection of individuals with social distance monitoring as a precautionary measure in reducing physical contact between people. The distance is going to be computed between the persons detected within the captured footage then compared to a hard and fixed pixels' values. The distance is measured between the central points and therefore the overlapping boundary between persons within the segmented tracking area. With the detection of unsafe distances between people, alerts or warnings can be issued to people for maintaining the safe distance.

A. *Numpy:*
Numpy is a python package which is used for scientific computation. It also provides matric, multidimensional array and it makes all operations on arrays very fast by assortment routines. It is highly optimized library for numerical transformation for example sorting, Fourier transformation, random simulation, linear algebra, I/O and much more [6].

B. *OpenCV:*
OpenCV (Open Source Computer Vision Library) is machine learning library for computer vision. It is open source and

contain more than 2500+ optimized algorithms of computer vision as well as machine learning. OpenCV also support multiple languages like C++, Python, Java etc. and has more than 18 million download and also has very big community of developers. OpenCV algorithms can be used identify objects, detect faces, do classification on human actions in videos, recognize text, track camera movement, extract 3D models, remove red eyes in images, make paintings using images, and acts as base for some high level technology like Virtual Reality (VR) and Augmented Reality (AR). [7] You can also increase performance by GPU (Graphics Processing Unit) acceleration. It allows us to use complex algorithm on our application without building them from scratch and allow us to focus on application. It supports all the major operating system like window, Linux , IOS etc. and with all of these functionality it is also very easy to use.

C. *Yolo:*

YOLO (You Only Look Once) is one of the popular algorithm which is used to detect object in and image but what makes it one of the best is that it gives high accuracy in detecting objects in videos at also it is very fast with of 23FPS (Frame per Second) which makes it usable for live video detection.This algorithm is used regression not classification and instead of selecting the interesting part of an Image, it predicts classes and bounding boxes for the whole image in **one run of the Algorithm.** It splits the images into multiple cells like 19x19 grid and then each cell predicts the bounding boxes.[8]

D. *Implementation:*

This model automatically detect extent to which distance protocols are being followed in particular area.

We can deploy this model on current surveillance system and drones used by police to monitor large area and make police work bit easy and utilize our police force efficiently by making tracking system happening in the area. It shows data in current time and give analytics of the area. It can also be used as alert system for police in case there are large population which violating the covid distance protocol.

We are using OpenCV to access camera and YOLO weight behind to detect people in camera view. Yolo is an algorithm which is able to give real time prediction to classify different objects with high frame (23 fps) which is needed to provide real time object detection in a live video and can also give analytics after processing.

We have used yolov3.weight and yolov3.cfg for detection purpose which is used by deep neural network to detect and processing.

Yolo can be used for any object detection because we are only want to detect humans we set it to only detect humans. After detecting person we check the confidence level and if it passed a certain threshold (>50%) then we used opencv to draw rectangles over them and then calculate distance between them using Euclidean distance algorithm. Where point will be distance points between two persons.

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Fig. 3: Euclidean Distance Formula [9]

Here we are using Euclidean Distance algorithm by providing the two people location to calculate distance between them.Now to we convert image to blob to give to algorithm to extract features then feed it to output layer. We have done classification of different types based on distance between two people.

Fig. 4: Colour classification based on distance between two humans

After doing classification we used this data to perform analysis of data and provide real time analysis and put them in screen to that this live feed data can be used and show data in below format. After the analysis we can check if lots of people are violating social distancing in certain area then we can efficiently utilize our police force in that area and increase security.

Fig. 5: Dashboard for live analysis

TABLE I
Depicting Comparison of Object Models

Model	TT (in sec.)	NoI	mAP	TL	FPS
Faster RCNN	9651	12135	0.969	0.02	3
SSD	2124	1200	0.691	0.22	10
YOLO v3	5659	7560	0.846	0.87	23

Here TT: Training time, NoI: Number of iterations, mAP: mean average precision, TL: Total Loss, FPS: Frame per second

From here we can conclude why YOLO is best for detecting giving best rate of Frame Per second (FPS) for detecting object model.

III. RESULTS

System development for this project has been completed based on Python 3.7, Open CV for image processing techniques and object distance detection model framework. Based on this developed system, research has been performed to test its Accuracy and results have been obtained. Specifically, the Bounding Pixel model has been used in this study as the key algorithm in person detection. For the program, the main video footage is captured from the entire scene set on the roads, where the camera is positioned high to gain overhead view. In this model we have also added a feature which ensures a timely e-mail is sent to the authorities alerting them that social distancing norms are being broken in this place. We have use ML, deep learning libraries to ensure that whenever more than 2 people are found in red box in the model (red box depicts people are present less than 2 feet distance from each other) an alert message is being sent to them to take appropriate action by the authorities.

Fig. 6 Depicting working of our model.

IV. FUTURE SCOPE

This project best describes social distancing we have to follow and is required to get protected from COVID-19. Lockdown can never be a permanent solution to cure spreading of COVID-19 thus this project helps in proper follow of guidelines about social distancing in market places. Improving our social distancing detector is to utilize a proper camera calibration [10].

V. CONCLUSION

Social distancing is one of the important precautions in reducing physical contact that may lead to the spread of coronavirus. Consequences of non-compliance with these guidelines will be causing the higher rates of virus transmission. A system has been developed using Python and Open CV library to implement two proposed features. **Social distancing** is especially **important** for people who are at higher risk for severe illness from COVID-19 The first feature is on detecting violations of social distancing, while the second feature is on detecting violations of entering restricted areas. Both features have been tested for accuracy. Based on the overall results, this study is seen to meet all of its objectives. However, there are some limitations to the results obtained. Based on the tests performed on the system, the results show that the object detection model used for detecting persons is having the difficulty in detecting people correctly in the outdoor environment as well as with the distant scenes. For further improvement in the future, a better object detection model can be implemented.

VI. ACKNOWLEDGMENT

We express our deep gratitude to Mrs. Gunjan Chugh, Assistant Professor, Department of Information Technology for her valuable guidance and suggestion throughout our research.

REFERENCES

[1] World Health Organisation. WHO Corona-Viruses Disease Dashboard. August 2020. Available online: https://covid19.who.int/table (accessed on 22 October 2020).
[2] WHO Generals and Directors Speeches. Opening Remarks at the Media Briefing on COVID-19; WHO Generals and Directors Speeches: Geneva, Switzerland, 2020.
[3] Olsen, S.J.; Chang, H.L.; Cheung, T.Y.Y.; Tang, A.F.Y.; Fisk, T.L.; Ooi, S.P.L.; Kuo, H.W.; Jiang, D.D.S.; Chen, K.T.; Lando, J.; et al. Transmission of the severe acute respiratory syndrome on aircraft. N. Engl. J. Med. 2003, 349, 2416–2422. [CrossRef]
[4] Ferguson, N.M.; Cummings, D.A.; Fraser, C.; Cajka, J.C.; Cooley, P.C.; Burke, D.S. Strategies for mitigating an influenza pandemic. Nature 2006, 442, 448–452. [CrossRef] [PubMed]
[5] Fong, M.W.; Gao, H.; Wong, J.Y.; Xiao, J.; Shiu, E.Y.; Ryu, S.; Cowling, B.J. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—Social distancing measures. Emerg. Infect. Dis. 2020, 26, 976. [CrossRef] [PubMed]
[6] https://numpy.org: depecting about numpy library.
[7] https://opencv.org/ : depecting about openCV library.
[8] https://towardsdatascience.com/yolo-you-only-look-once-3dbdbb608ec4
[9]https://www.sciencedirect.com/topics/mathematics/Euclideandistance#:~:text=Only%20Euclidean %20distances%20may%20be,a%20h %20%3D%20m%20s%20h%20.
[10] https://www.pyimagesearch.com/2020/06/01/opencv-social-distancing-detector/.

Fake News And Rumor Detection Techniques

Dhruv Rustagi[1], Harshit Jain[2], Gaurav Maindola[3], Monica Batra[4]

[1,2,3]*Student B-Tech-IT, ADGITM, GGSIPU University, New Delhi, India*

[1]rustagidhruv13@gmail.com

[2]harshitjain137@gmail.com

[3]maindolagaurav23@gmail.com

Assistant Professor IT Engineering, ADGITM, GGSIPU University, New Delhi, India

[4]monica.batra88@gmail.com

Abstract **- Fake news and rumors represent the foremost popular sorts of false and unverified information respectively, and will be detected as soon as possible for avoiding their dramatic effects. Interest rates on successful acquisition strategies have been growing steadily over the years. In this paper we examine various ways in which automatic discovery of false stories and rumors has been raised within recent publications. Basically, we focus on the main points. First, we report and discuss the various interpretations of false stories and rumors that are considered within the literature second, we highlight at the collection of relevance of relevant data for non-discriminatory disclosure is problematic and that we present various, accepted methods of collecting this information and also because data sets are publicly available and why they are being distributed. Third, We explain the factors that are considered in matters other than the means to obtain rumour. Fourth, we provide a comprehensive analysis of the various techniques used to make rumour and final false stories**

I. INTRODUCTION

False stories or jokes or fake news, or fraudulent stories are all kinds of stories that contain deliberate information or deceptions that are spread through traditional media or online forums such as WhatsApp, Facebook, Instagram, Twitter. It's a Big Data problem. We are trying to resolve it with small amounts of data. [17] That, in short, is the three main points of our paper. We will no longer read an unfamiliar area covered with many recent papers, reports and news articles about how the spread of false news, and that mistrust of news is a democratic problem. [18,23] In general, the importance of false stories has increased in the politics behind the truth. In media stores, the ability to attract viewers to their websites is needed to generate revenue for online advertising. Publishing an article with false content that attracts advertisers for the benefits of users and improves ratings. [31] The easy availability of online advertising revenue, increased political divisions and the popularity of social media, especially Facebook, News Feed, have all contributed to the spread of fake, competing news and official news.

Rumors sometimes may spread very quickly over social media platforms, and rumor detection has gained great interest in both academia and industry recently. [32] Government authorities and social media platforms are also taking efforts to defeat the negative impacts of rumors.

II. HOW FAKE NEWS AND RUMORS GETS VIRAL ON SOCIALMEDIA?

Fake News spreads very fast throughout the nation as the news is somewhat interesting and helps to increase their own publicity. The major reasons how the fake news gets viral on social media are:

A. NO Gatekeepers:

Anyone can access and share any technology on Internet by posting anything on various social medias like Facebook, Twitter since they are free house.[20]

B. Speed and Information Overload:
Anything that is posted or shared spreads in a very fast speed before anyone can check it. And by time it may lead to danger

C. Permanence of Postings:

Once you have posted or shared any image, video, audio etc. on social media, it's there forever and can't be removed or deleted for everyone. All the news is shared without getting the proper context. Efforts to suppress media often result in them gaining further prominence, as "The Internet routes around censorship." [34,35]

D. Search Engine Optimization:

Search engine optimization, or, SEO, is strategies that website owners use to attempt to have their websites show up higher in google search results. These range from perfectly legitimate SEO practices like making sure every page has a title, structuring information well, and Keeping content up to date [38]

Fig 1 : Search Engine Optimization

III. WHY FAKE NEWS AND RUMORS ARE SPREAD?

In the last few years the fake news and rumours are shared and posted on the social media in a very fast manner. It is done to gain their publicity and get to have their interest on everyone's mind. Some people share this news for their satire or to yield a joke or it may be April Fools [37] day related. Fake news and rumours are Spread throughout the public

A. Make a Statement

B. Generate Toxicity

C. Gain Followers

D. Make Money

E. Influence Politics

F. Garner public opinion for hatred.

IV. DETECTION METHODS OF FAKENEWS

A. Text-Based Fake News Detection:

Fake News can be detected using the text-based approach only when no previous information of the subject domain is available while assuming that faux news spreads faster than true data. Using the rumour dataset 126000messages which was tested on three million people. It was checked that what kind of messages were easier to discriminate between false & True. The Same type of Fake information is spread using multiple online accounts to pass it off as Public opinion [12]. A more practical manner of achieving this by exploitation social bot nets-that retweet and share constant messages indiscriminately bring home the bacon quality and achieve bigger unfold and coverage[14]

B. Text Sentiment Analysis:

Emotional analysis is also known as the digging of ideas seeks to understand the practical meaning of sentences and phrases. Provides levels of classification in publications made by authors also called polarity. It can be as simple as a binary level like good and sometimes neutral. The same tools and methods were used by [15] who used a weakly supervised measured and random step to create dictionaries and a word abg. Similarly [16] using moving an average textual comment over a period of time, it has been found that negative and positive emotions expressed by users on twitter are real indicators of voter confidence and approval ratings of the president. While the emotional analysis from the text goes beyond polarity it can include the determination of the emotional state of the authors such as anger, anxiety, depression and happiness. [25,39]

V. TEXT CLASSIFICATION FOR FAKENEWS

An intuitive framing of the fake news problem in NLP would be to ask how we can classify news text into fake and legitimate instances. This is especially true in the case of full text - unlike tweets or articles that are being circulated on social media - because text editing is highly dependent on the language features of long text.

A. NLP-based approaches:

A feature-based approach, which involves the extraction and analysis of language indicators to identify specific targeted objects (e.g. reviews of counterfeit products from real ones) have become a very powerful model with interpretive results.

B. In-depth learning models:

Deep learning has taken a lot of NLP functions but usually in areas where great training details are available. In text classification, Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs) and Attention models competed with content-based models.[13] For more classification of misinformation of fake news and rumors, refer figure 2.

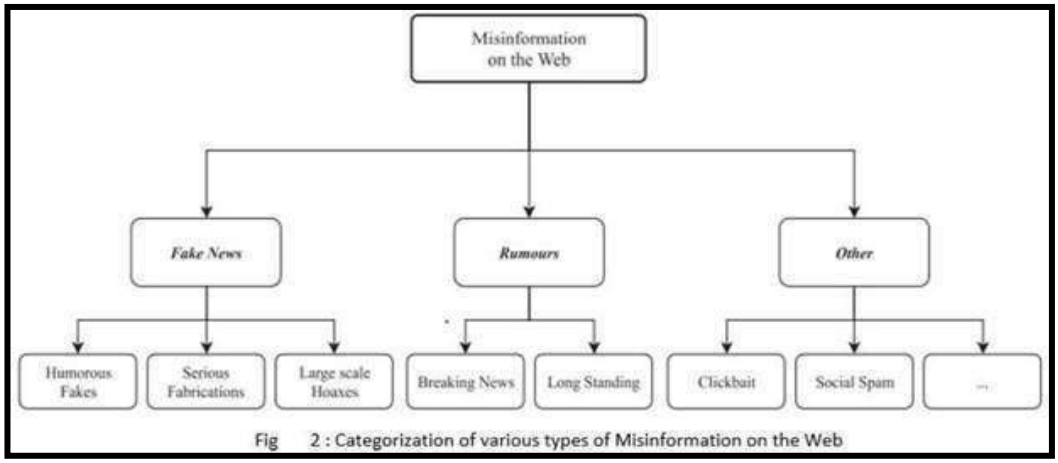


Fig 2 : Categorization of various types of Misinformation on the Web

VI. APPROACHES TO THE FAUX NEWSDRAWBACK

A. Educating the Public:

Educating efforts could also be increased, beginning at the varsity level, with media acquisition; and a general education towards empowering an accountable citizenship raised in civil and democratic values, who is able to perceive the competitive pressures of capitalist societies, as well as the influence of lobby teams, political parties. [4]

B. Analyzing and Curtailing the spread:

Fake news unfold quick. It spreads quicker and penetrate social network to a much bigger extent than credible news. This might be due to its novelty, its capability to urge outrage, or its role in conforming the preceding biases of the reader. The novelty and outrage might justify why Fakebook's effort of flagging debunked faux stories backfired [1,8].A locality of the matter reside in echo chambers or filter bubbles, that suggest that some individuals area unit exposed to a just one purpose of read, and may be easier to believe stories that return that point of read.

C. Automatic Checking:

There area unit clear benefits to playing offer automatically; it should be do neat scale, and it saves moderators from having to type through at the simple stun pleasant content. This sort of automatic checking is regarding the content and claims among the story itself, not regarding data like supply or rate of unfold. Procedure fact checking makes an attempt to find makes an attempt claims throughout a story or rumour, and check them against reliable sources.[5]

D. Manual Checking:

Manual checking of false statements, rumours and fake news articles online play a vital role in containing the spread. Two Broad Classes of efforts can be identified: using fact-checking websites, and performing manual checking on specific social media sites [40] Fact-checking websites provide verification of claims that they find, or that users submit. They have the advantage of using qualified journalist and other professionals, who are able to research and verify claims and news stories.

VII. APPROACHES BASED ON PROPAGATION PATH ANDNETWORK

Rumors spread through social media within the form of shares and re-shares of the source post and shared posts, leading to a diffusion cascade or tree. [3,7] The trail of re-shares and other propagation dynamics are utilized for rumor detection. We group current studies into 2 types:

A. Propagation-based:

When using cascades to detect rumours, one either distinguishes them by computing the similarity of fits cascade to it of other true/false rumours, or by generating a cascade representation that facilitates distinguishing false and true rumours

B. Network-based:

Network-Based rumor detection constructs flexible networks to indirectly capture rumor propagation information. The constructed networks can be homogeneous, heterogeneous, or hierarchical. Construct a network consisting of users, messages and events, using a PageRank-like algorithm.[21]

VIII. LOCAL WAY TO CLARIFY THE FAKENEWS

A. *Check the Date*

B. *Check the Byline*

C. *Evaluate the Tone*

D. *Reverse Image Search*

E. *Reverse Video Search*

F. *Check the URL*

G. *Evaluate Crazy Claims*

H. *Google Search on the page you're not sure about.*

IX. CASESTUDIES

A. PulwamaAttack,2019

India's social media misinformation downside is spilling out on to the streets ,,as a result of the world's largest market place for Facebook and WhatsApp with a combined base of 500 million struggles to remained on the unfold of pretend news[10].Yet it's the set terribly platforms that were accustomed flow in to false hoods that area unit currently being mobilized to administer associate degree remedy. On Valentine Day, Asian Country witnessed the deadliest ever attack on its troops once forty-four paramilitary cops were killed once a terrorist crashed an automobile laden with explosives in to military convoy[30].The attack ,that additionally left seventy out of action, occurred in Indian-administered geographical area, the region at that center of a seven-decade recent feud with neighbor west Pakistan[11].It Elicited widespread anger in Asian Country, with several line of work for retaliate action against the rival country for providing refuge to the phobia cluster caleed Jaish-e-Mohammad that claimed responsibility for the attack. Mostly the fake news and rumours are speeded on media [12,36]

B. False Rumors on Coronavirus could cost lives, say Researchers,2020

The Coronavirus outbreak has sparked what the global health organization is looking on "infodemic"-an awesome amount of data on social media and websites. The False statements range from a conspiracy theory that the virus may be a manmade bioweapon on to the claim that quite 100,000 have died from the disease(as of on ,the amount of reported

fatalities reported at 2200plus)[22].In early January ,a few weeks after china reported serious cases, a Un Company launched screening program to ensure the facts about the newly identified virus were made public. The project is called the short EPI-WIN of the WHO information [28,29] Professor Paul Hunter, a specialist in new corona virus infection, now called covid-19 and Dr.Juli Brainard from UEA's Norwich Medical School, said efforts to Disseminate relevant information on social Media and correct false news could save lives[37]

X. DISCUSSIONS

Facebook's new policy could be a sensible commencement, however all folks have a task to play, as a result of pretend news affects everybody. [26,27]

A. At a minimum, you must avoid spreading pretend news, that is sort of as dangerous as making it. Nobody needs to be referred to as a "liar," however if you unfold pretend news while not creating a minimum of a bottom effort to validate it, that's just about what you're doing.

B. You can additionally report it mistreatment Facebook's tool if that's wherever you discover it or by commenting that the story is pretend so others WHO see it'll additionally see your correction, ensure you embody a link to any truth checking supply so folks will see why you suspect it's pretended.[33]

C. You can educate your friends and family and particularly youngsters concerning the presence and danger of FauxNews. Mention it at the dining table, share the following pointers and raise others if they apprehend of samples of pretend news or "urban myths".

D. Step in after you see an addict sharing a pretend story on social media. Embody a link to a post on Snopes or Alternative proof of that refutes the story [6]. Thus, Folks are going to be discouraged from connecting or, worse sharing to pretend story

E. Fake News Intervention. Fake news studies, e.g., Laser et al. [2018], [2] have emphasized the importance of business models adopted by social media sites to address fake news intervention, which suggests shifting the emphasis from maximizing user engagement to that on increasing information quality, e.g., using self- or government regulations [9]

XI. CONCLUSION

We have discussed the different approaches to the problem of Fake news and Misinformation, some of them relating to how to educate the public or to how to stop the spread of such premicoious news. We focus on addressing the issue as a text-sharing problem, that is trying to automatically determine whether a particular news article is false. By 'fake' we mean an article containing unconfirmed or un true claims, or attempts to spread in accurate information. In order to make automated Classification of Media, Modern NLP methods and machine learning methods require a large amount of training data. As Computer Language Researchers, we feel however, that we cannot determine for ourselves which stories are not true. Although significant improvement has been made in eliminating rumours on social media, there are still many challenges to overcome. Based on the review of previous studies and also our experience in both research and practical system implementation of rumour detection, here we present several directions for future rumour detection research.

REFERENCES

[1] Virginia Woolf, Three Guineas, originally published in 1938, reprinted in Lapham's Quarterly, A History of Fake News(2018).
[2] Cited in Lazer et al.(2017).
[3] Ref: ifla organization- "How to spot fakenews."
[4] ShorensteincenterOrganisation on Media, Politics and Public Policy free course identifyingmisinformation
[5] CredibilitycoalitionOrganisation
[6] S. Wu, J. M. Hoffman, W. A. Mason, and D. J. Watts, "Who says what to whom on twitter," in Proceedings of the 20th international conference

on World wide web. ACM, 2011, pp.705–714.

[7] J. W. Pennebaker and L. A. King, “Linguistic styles: Language use as an individual difference.” vol. 77, no. 6, p. 1296,1999.

[8] M.L.Newman,J.W.Pennebaker,D.S.Berry, andM.Richards,“Lyingwords:Predictingdeceptionfromlinguistic styles,” Personality and social psychology bulletin, vol. 29, no. 5, pp. 665–675,2003.

[9] “Fake news - political scandal words.” [Online]. Political scandal words andrumours.

[10] O. Ajao, D. Bhowmik, and S. Zargari, “Fake news identification on twitter with hybrid CNN and RNN models,” in 9th Int’l Conference on SocialMedia.

[11] Juan Cao, Junbo Guo, Xirong Li, ZhiweiJin, Han Guo, and Jintao Li, Automatic Rumor Detection on Microblogs, 2018,Carlos Castillo, M. Mendoza, and B.Poblete.

[12] Information credibility on twitter. WWW 2011. Cheng Chang, Yihong Zhang, Claudia Szabo, and Quan Z. Sheng. 2016.

[13] Extreme user and political rumor detection on twitter. ADMA 2016. WeilingChen, Chai Kiat Yeo, Chiew Tong Lau, and Bu Sung Lee.2016.

[14] Behavior deviation: An anomaly detection view of rumor preemption. 2016 IEEE IEMCON Alton Y. K. Chua and Snehasish Banerjee. 2016. Linguistic predictors of rumor veracity on the internet. 2016MCECS. L. Derczynski, K. Bontcheva, M. Liakata, R. Procter, G. W. S. Hoi, and A. Zubiaga, Semeval-2017 task 8:Rumoreval:

[15] Determining rumor veracity and support for rumors, SemEval 2017 Omar Enayet and Samhaa R El-Beltagy. 2017. Niletmrg at semeval-2017 task 8: Determining rumor and veracity support for rumors on twitter.SemEval17.

[16] Bartolotta, Devin (December 9, 2016), "Hillary Clinton Warns About Hoax News On Social Media", WJZ-TV, retrieved December 11,2016.

[17] Wemple,Erik(December8,2016),"Facebook'sSherylSandbergsayspeopledon'twant'hoax'news.Really?",The Washington Post, retrieved December 11,2016.

[18] Tufekci, Zeynep (January 16, 2018). "It's the (Democracy-Poisoning) Golden Age of Free Speech".Wired.

[19] Wikipedia aboutfakenews

[20] Vosoughi, Soroush. “THE SPREAD OF TRUE AND FALSE NEWS ONLINE”, MIT Digital Retrieved March 5, 2019.

[21] Itti, Laurent (2005) “ Bayesian Surprise Attracts Human Attention” Retrieved March 6,2019.

[22] NBC news article about german reporter stripped cnn journalist year awards fabricatingstories

[23] N. Newman, W.H. Dutton, G. Blank, Social media in the changing ecology of news: the fourth and fifth estates in Britain, Int. J. Internet Sci. 7 (1) (2012) 6–22.

[24] A. Hermida, Twittering the news: the emergence of ambient journalism, Journal. Pract. 4 (3) (2010)297–308.

[25] S. Vieweg, Micro blogged contributions to the emergency arena: discovery, interpretation and implications, in: Proceedings of the ACM Conference on Computer Supported Cooperative Work, ACM, 2010, pp.241–250.

[26] C. Kang, A. Goldman, In Washington Pizzeria Attack, Fake News Brought Real Guns 5(2016).

[27] J. Ma, W. Gao, P. Mitra, S. Kwon, B.J. Jansen, K.-F. Wong, M. Cha, Detecting rumors from microblogs with recurrent neural networks, in: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI, address, 2016, pp. 3818–3824.

[28] V.L. Rubin, Y. Chen, N.J. Conroy, Deception detection for news: three types of fakes 52 (1) (2015)1–4.

[29] WhatsApp rumours have led to 30 deaths in India. In this social media disinformation age, the only question is: who’s next? Kunal purohit(2019).

[30] BBC news on social media rumours in india: counting thedead

[31] F. Wang, M. Lan, Y. Wu, ECNU at SemEval task 8: Rumor evaluation using effective features and supervised ensemble models, In Proceedings of the Eleventh International Workshop on Semantic Evaluation(SemEval),2017, pp.491–496.

[32] S. Wang, I. Moise, D. Helbing, T. Terano, Early signals of trending rumor event in streaming social media, in: Proceedings of the IEEE Forty First Annual Computer Software and Applications Conference (COMPSAC), 2, 2017, pp.654–659.

[33] S. Wang, T. Tereno, Detecting rumor patterns in streaming social media, in: Proceedings of the IEEE International Conference on Big Data,2015.

[34] W.Y. Wang, “Liar, liar pants fire”: a new benchmark dataset for fake news detection, in: Proceedings of the Fifty Fifth Annual Meeting of the Association for Computation Linguistics (Volume 2: Short Papers), Association for Computational Linguistics, 2017, pp.422-426.

[35] G.Widmer,M.Kubat,learninginthepresenceofconceptdriftandhiddencontexts,Mach.Learn.23(1)(1996)69– 101.

[36] Buzzfeed news article whatsapp fake news tip line Indian election nothelpline.

[37] Theguardian article about coronavirus false rumours cost lives sayreseachers.

[38] Search engine optimization, Google search central help community, Googledevelopers.

[39] The effect of social media on elections: Evidence from the united states, Thomas fujiwara, Karsten Muller, Carlo schwarz April2021

[40] Bigdataandqualitydataforfakenewsandmisinformationdetection,Bigdata&society,Fatemehtorabiasr,Maite Taboada.

A mm-scale Wireless Energy Harvesting Sensor Platform

Shivaansh Parmar[1], Akshita Sharma[2], Shimpy Goyal[3]
[1,2]*Student, Dept. Of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology and Management Bazpuri, India*
Indore, India
[1]akshita.2507.sharma@gmail.com
[2] parmarshivaansh@gmail.com

[3]*Assistant Professor, Dept. Of Information Technology, Dr Akhilesh Das Gupta Institute of Technology and Management New Delhi, India*
shimpygoyaluiet@gmail.com

***Abstract*- We explore the critical role that power plays in the development of a mm-scale sensor. We argue that any practical deployment of a mm scale system must have a significant energy harvesting system. We demo a 1.0 mm^3 general-purpose sensor node platform with heterogeneous multi-layer structure capable of short-range (order cm) wireless transmission. The sensor platforms gets benefitted by dividing the system into varying degrees of components and hide the complexity of each part behind an abstraction and interface by allowing the addition/removal of IC layers. A self-adapting power management unit is proposed for efficient battery voltage down conversion for a wide range of battery voltages and load current. The power management unit also adapts itself by monitoring energy harvesting conditions and harvesting sources and is capable of harvesting from solar, thermal and microbial fuel cells**
Keywords - Ultra-low power, wireless sensor node, smart dust.

I. INTRODUCTION

Continuous evolution of fabrication technology has led to the integration of computational capabilities of computers in an increasingly smaller volume. These Fabrication systems are used to create very small yet highly capable systems, as well as new multi-core and/or networking technologies that push the upper limits of modern computing performance. This, in turn has led to the development of various computing platforms, ranging from large computers that occupied large space equal to a room to portable handheld devices. According to Bell's Law, roughly every decade a new, lower priced computer class forms based on a new programming platform, network, and interface resulting in new usage and the establishment of a new industry. These computers are developed by using fewer components or fractional parts of state-of-the –art computing systems. In addition, the size of each subsequent class of computing systems becomes approximately 100 times smaller than its predecessor. Looking at the History of computers and comparing them we find that the first computers introduced in the 1940s were room sized or even as large as a small building. After that smaller and more affordable computers were introduced in the form of workstations in 1970s and then the personal computers came in 1980s, and mobility was added to the computers with laptops in 1990s and portable handheld devices in 2000s. Recent researches have shown that there is a potential of wireless sensor nodes, which is expected to be the next generation of the computers after handheld portable devices.
These wireless sensor nodes can be used in wide range of applications including health monitoring, environmental monitoring, biomedical implants, and creating smart buildings.

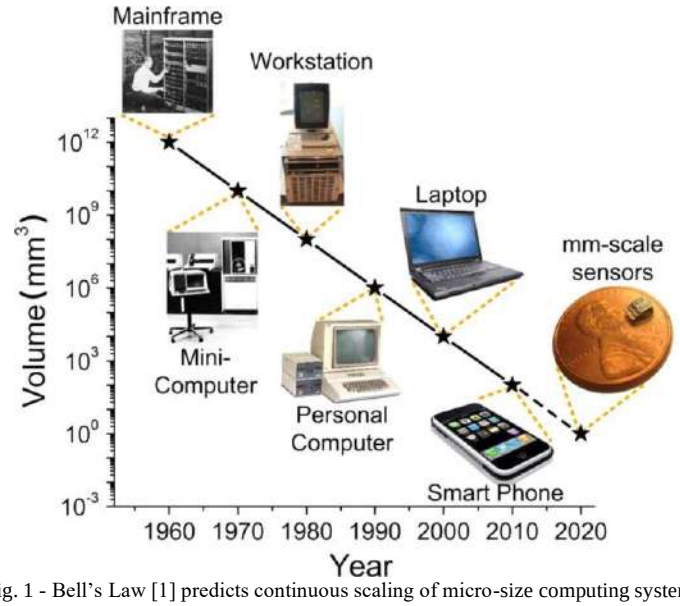


Fig. 1 - Bell's Law [1] predicts continuous scaling of micro-size computing systems.

The WSN is built of "nodes" – from a few to several hundreds or even thousands, where each node is connected to one (or sometimes several) sensors. Each such sensor network node typically has several parts: a radio transceiver with an internal antenna or connection to an external antenna, a microcontroller, an electronic circuit for interfacing with the sensors and an energy source, usually a battery or an embedded form of energy harvesting. A sensor node might vary in size from that of a shoebox down to the size of a grain of dust. Bulky batteries are included to power the circuit components with adequate lifetimes. The result is milliwatt-level systems that are centimeters or tens of centimeters on a side, failing to realize the vision of "smart dust". Smart dust requires mm^3-scale, wireless sensor nodes with perpetual energy harvesting Technologies to the effective realization of sub-centimeter computing.

Nearly all of the design challenges in the creation of a "dust-scale" sensor network system can ultimately be traced back to one common root: power. Storing power is expensive in terms of volume, but to provide an effective wireless sensor node, several power-hungry activities must be supported: a variety of sensing modalities, computation for at least basic data processing, and communication to share sensed data. To achieve long lifetime in a mm -scale system, power consumption of each component must be aggressively reduced to stay within a stringent power budget of 10 nW.

To this end, we propose a 1.0 mm^3 general purpose reconfigurable sensor node platform with a heterogeneous stackable multi-layer structure. The key components implemented to realize this form-factor includes ultra-low power IC (Inter-Integrated Circuit), a 228 pW standby power optical wakeup receiver, ultra-low power power management unit (PMU) and brown-out detector (BOD).

II. SYSTEM OVERVIEW

The 1.0 mm^3 sensing platform is designed with stacked integrated circuit (IC) dies fabricated in three different technologies. Fig. 2 shows the dimension of each die and the wire bonding scheme for electrical connectivity of the sensor system. To enforce 1.0 mm^3 volume, each layer measures less than 2.21× 1.1 mm and the length of each layer has to be reduced by 140µm compared to the lower layer to provide enough clearance for bond-wires. The height of each IC layer is thinned to 50 m, while the custom thin-film Li battery is 150 μm thick. The system's die-stacked structure with wire bonding provides maximum functionality (or

silicon area) per unit volume and also enables easy expansion of the system with additional layers. End users can create a sensor system for new applications by designing an application-specific layer in a preferred technology, which complies with the system power and energy budget, and providing an identical inter-layer communication interface.

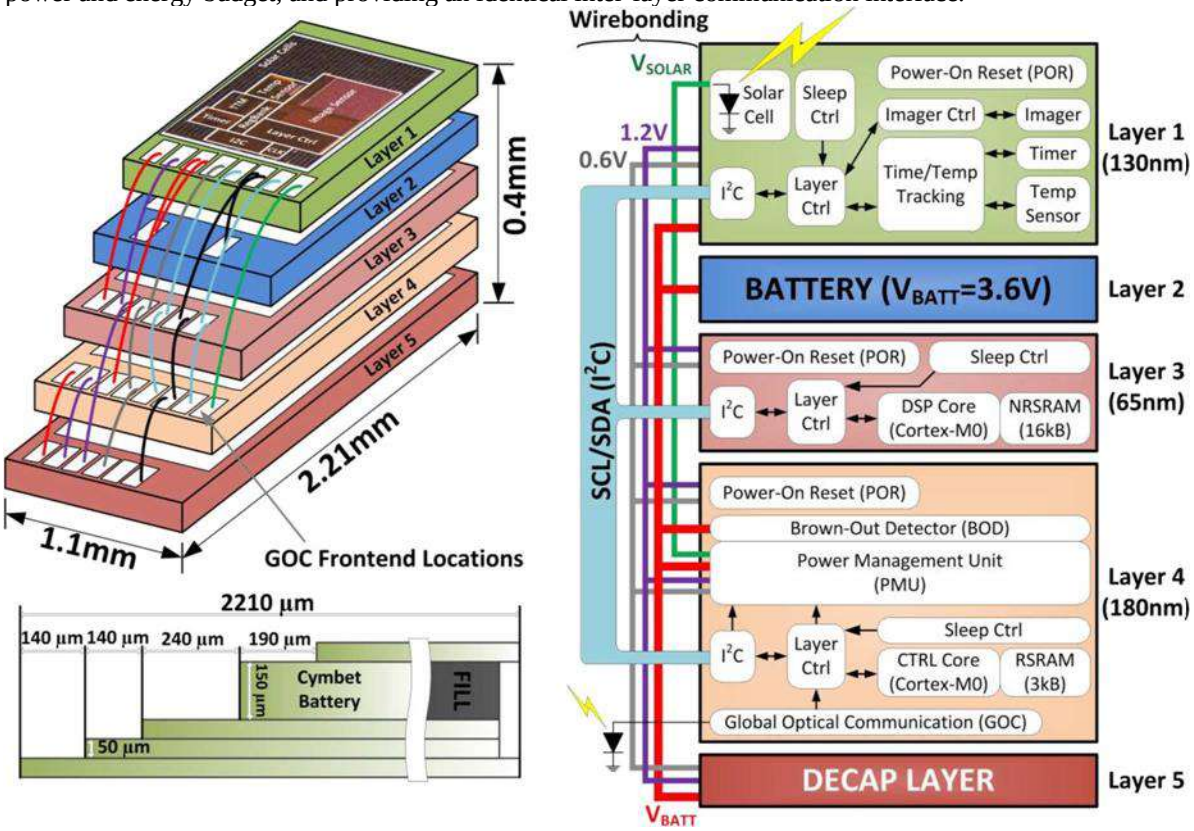


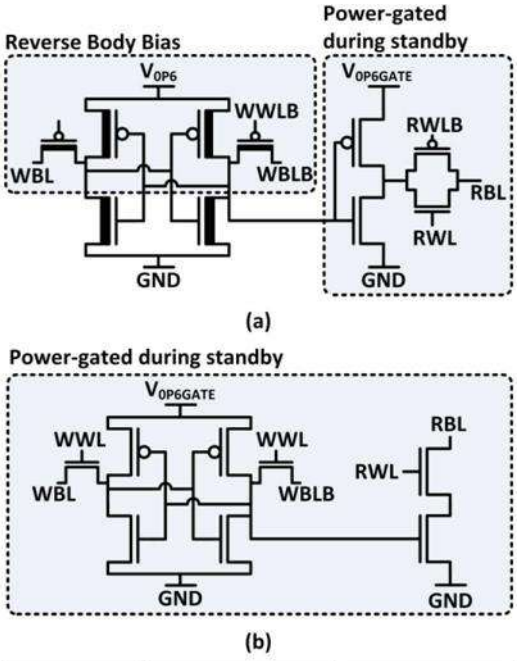
Fig. 2 shows the system block diagram. The various components in the system are categorized as CPU, memory, power management, timer, and sensors.

A. CPU

The operations of dust-scale sensor system are sequenced and managed by a control microprocessor, which requires low computational performance and hence can be rearranged to improve the efficiency of retrieval or processing for lower power operations. On the other hand, there are some sensors like imager which requires high performance for digital signal processing (DSP) operations. Because of this reason, two ARM® Cortex-M0 processors are located in separated layers with different functionality as follows:

1) 1. The DSP CPU effectively handles the operations of imager sensor (or other sensors) and is thus built in 65nm CMOS (layer 3) with a large 16KB non-retentive SRAM(NRSRAM). Because of such an advance technology node the DSP CPU runs faster than control CPU (fabricated in 180nm) and has larger memory capacity which is required to perform complex DSP operations. Due to high leakage current in this process the power supplied to some of the components of SRAM has to cut down during standby mode and hence the SRAM is non-retentive.

2) On the other hand, control CPU manages the system operations by using 3kB retentive SRAM(RSRAM) which is always in ON mode to maintain the stored operating programs effectively. It is built in low leakage 180nm CMOS (Layer 4).



	RSRAM	NRSRAM
Topology	Low leakage 10T[10]	Low voltage 8T[11]
Bitcell Area	17.48 μm^2	1.548 μm^2
Standby Power	3.3 fW/bit	14 pW/bit

Fig. 3 SRAM topologies used for (a) NRSRAM, (b) RSRAM, and their comparison.

B. Memory

RSRAM, For RSRAM,a10T bitcell with high threshold voltage (V_{th}) I/O transistors[11] issued to minimize the leakage power. Although it exhibits very low leakage power (3.3 fW/bit), the large bitcell area of 17.48µ m limits its capacity on Layer 4 to 3 kB due to the large spacing requirements of I/O transistors and older process technology. For NRSRAM, larger capacity is achieved with an 8T SRAM structure [12] with regularV_{th}transistors, and bitcell area of 1.548 m . However, the low V_{th} of regular transistors in 65 nm technology compared to I/O transistors in 180 nm incurs a high bitcell leakage power of 14 pW/bit, which necessitates the power-gating of the entire NRSRAM array during standby. This means that execution code and data in NRSRAM is lost during standby; therefore, execution code is first transferred from RSRAM whenever DSP operation is activated and the DSP processor functions much like a co-processor.

a. Power Management and Harvesting

For energy harvesting solar cells and a lower power imager are placed at the top of the sensor on the top layer (Layer 1) for light exposure. Power consumption ranges from 11nW in sleep mode up to ~40µW in active mode. A flexible PMU allows harvesting for perpetual operation from multiple sources, including solar, TEG (Thermo-Electric Generator), and microbial fuel cells. A brown out detector (BOD) monitors the battery voltage and when brown out is detected, it shuts down entire system to prevent Li battery damage due to excessive drainage.

b. Timer

A gate-leakage-based timer [10] and temperature sensor is also implemented in Layer 1, which is fabricated in 130 nm CMOS for gate-leakage current optimization for timer accuracy. Time tracking with temperature compensation is implemented using this timer, providing a timing reference to synchronize the radios that are included in our second system configuration.

c. Sensors

A low power temperature sensor and 96× 96 pixel low power CMOS image sensor are implemented in Layer 1.

III. LOW POWER I²C COMMUNICATION

The die-stacked structure of the proposed sensor platform requires communication among different layers. Due to pad count limitations arising from the 1 mm^3 form factor, the number of wires used for communication is of critical concern. I²C [13] is a widely used industry standard serial communication protocol that only requires two wires—serial clock (SCL) and serial data (SDA)—and is easy to expand with any I²C-compatible devices. However, conventional I²C relies on pull-up resistors, as shown in Fig. 4, which consume mW-order power when the wires are pulled down. Assuming a 1.2 V supply voltage and 1 kΩ pull-up resistance, the average pull-up current for both wires is 1.2 mA, which results in 1.44 mW power wasted simply for pull-up current without considering decoder and driver overhead. This is clearly unacceptable for a sensor platform targeting tens of μW active power. Therefore, a modified communication protocol is required to meet the stringent power budget of the system, while maintaining compatibility to the standard I²C protocol to enable expansion with C compatible devices.

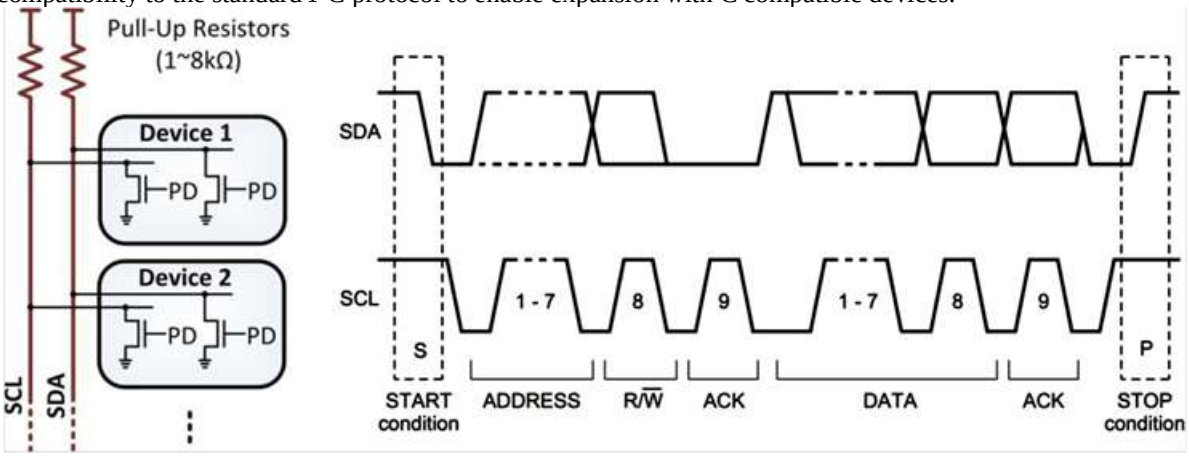


Fig.4. Conventional I²C circuit diagram and data transfer waveform.

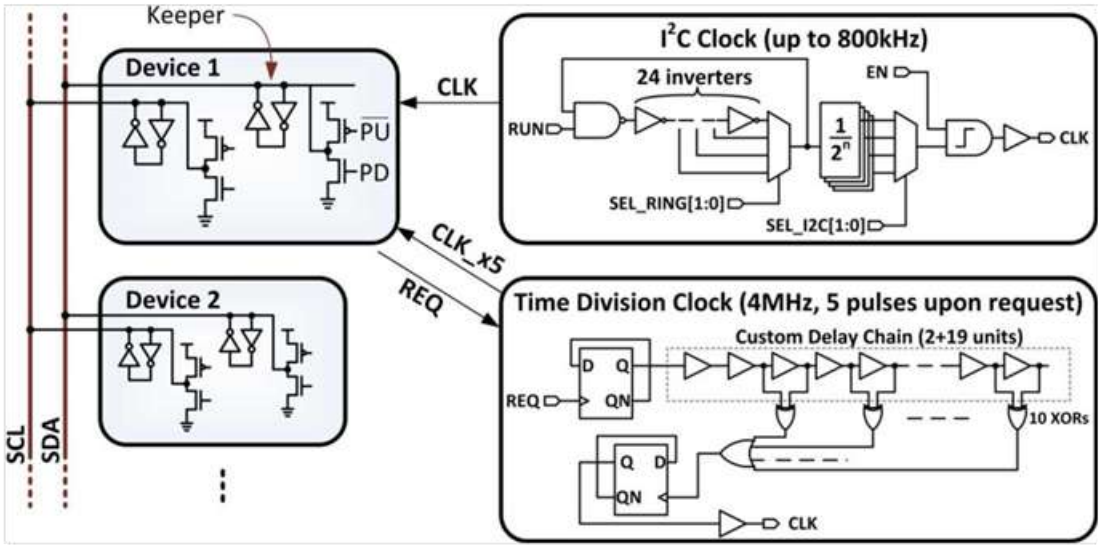


Fig.5. Proposed low power I² C circuit diagram.

Fig. 5 shows the circuit diagram for the proposed low power I^2C protocol. Pull-up resistors in conventional I^2C act as 1) a pull-up device when the wire is at ground potential and pull down is released by attached devices, and 2) a keeper for holding high once the wire is fully charged to the supply voltage. To provide the pull-up device function without use of the pull-up resistor, the SCL-low cycle is divided into five sub-cycles where a master device always pulls up SDA in the second sub-cycle and holds high using small keepers. Any attached device can pull-down in the fourth sub-cycle, which complies with the I^2C standard—SDA can change only when SCL is low and a layer pulling down has the higher priority. The length of the sub-cycle is determined by a sub-cycle clock generator (Fig. 5) and marginal sub-cycles (first, third, fifth) provide margins for die-to-die sub-cycle length variations, which are directly related to the variation in the sub-cycle clock generating circuit. With provided margins, sub-cycle length variation up to 25% can be tolerated.

To provide the keeper function of the removed pull-up resistor, a keeper is attached to each wire. The proposed low power I^2C scheme allows communication between a low power I^2C master device and standard I^2C slave devices if a proper power supply is provided. The only additional cost for such a configuration is occasional short circuit current during the second sub cycle, when the low power I^2C master device pulls up SDA and the standard I^2C slave device pulls down SDA for acknowledge. Energy overhead due to such short circuit current in I^2C write to slave would not be significant since it occurs only during one sub-cycle and only for acknowledge operation, which is once every 8 bit transmission. However, the overhead is larger for read operations between low power I^2C master and standard C slave. In this case, the master should pull SDA high for each data cycle due to the lack of a pull up resistor, which can generate short circuit current whenever the transferred bit is a 0. Even so, current consumption in this situation is still lower than standard I2C system which employs pull up resistor that continually draws current when SDA or SCL are pulled low. To tolerate a current surge, external power supplies with higher current capacity may be required, which could be provided by simply sharing the power supply for the attached standard I^2C device.

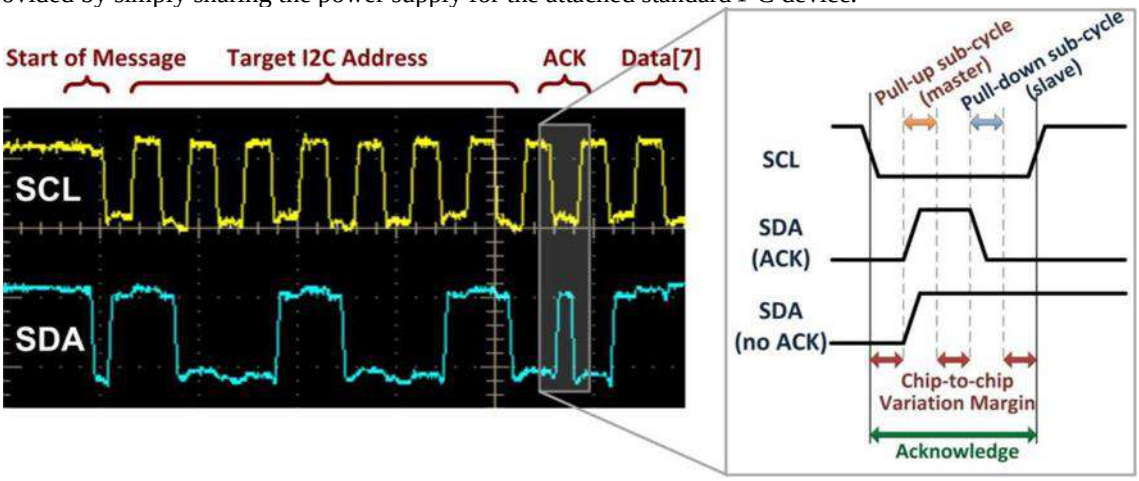


Fig. 6. Measured I^2C waveform and illustration of SCL-low sub-cycle

Fig. 6 shows the measured low power I^2C waveform. The figure clearly shows that in each SCL-low cycle, SDA is raised only in the second sub-cycle and pulled down in the fourth sub-cycle, which is clearest in the acknowledge cycle where both pull-up and pull-down operation are performed. Measured energy consumption is 88 pJ/bit, which is more than an order of magnitude lower than 3.6 nJ/bit, the theoretical minimum energy needed to drive the wires in standard I^2C protocol (excluding overhead) for decoding and driving logic at the maximum data-rate in 'fast mode' I^2C, 400 kbps [13].

IV. POWER MANAGEMENT UNIT (PMU)

The thin-film Li-ion battery in the proposed sensor platform outputs voltages as high as 4.1 V while low power electronics often operate below 0.5 V, resulting in a challenging DC-DC conversion ratio. The PMU must enable both up-conversion during harvesting and down-conversion in the absence of harvesting conditions. Further, it must accommodate a 1000 spread in current draw between sleep/active modes and low/high harvesting conditions. Having switches that can accommodate 10 s of µA in active mode, while at the same time maintaining high efficiency conversion with load currents as low as single digit nAs in sleep mode, is extremely challenging. Finally, different harvesting sources have varying optimal operating points that change with harvesting

conditions. For instance, a solar cell has an open circuit voltage as low as 350 mV in indoor conditions, increasing to 530 mV in sunlight, while a microbial fuel cell (MFC) has an open circuit voltage that tends to fluctuate over time from 500 to 800 mV. These requirements call for the PMU in the proposed sensor platform to be highly adaptive to load current and harvesting source conditions.

The proposed PMU is fully integrated with a chip area of 0.95 mm^2 (Active region: 0.24 mm; MIM-capacitance area: 0.71 mm) and automatically adapts to different harvesting conditions and can be configured to operate with different harvesting sources. The PMU consists of two ladder-type switch capacitor networks (SCNs), as shown in Fig. 7; one for converting between battery and processor voltages (*battery SCN*) and one for converting between processor and harvester voltages (*harvester SCN*). During down conversion, when the processor draws current from the battery, the battery SCN is automatically reconfigured between 5× and 6× modes, providing the ability to adapt to different load currents and battery condition from 3–4.1 V.

During harvesting, both SCNs up convert the harvested voltage and provide current to the processor and battery. The harvester SCN has 2 and 3 modes and is connected to the battery SCN in one of three possible configurations to adapt to different harvesting sources and harvesting conditions. The key challenge in efficient harvesting is to determine the configuration that forces a voltage at the harvester leading to extraction of maximum energy from the harvester. Measurement with a solar cell in Layer 1 showed that the extracted power peaks when the extraction voltage is approximately 0.83 of the open circuit voltage (V_{OC}). However, the varies between 350 and 500 mV depending on light conditions, and hence the harvest extraction voltage must be adjusted accordingly. In addition, different energy sources have different power profiles leading to the need to adjust the fraction of V_{OC} at which the harvested power is extracted.

The proposed PMU adapts to different harvesting conditions using a two-phase process. First, the battery and harvester SCNs are disconnected for a short monitoring time period. During this time, the harvester develops its open-circuit voltage, which is up converted by the harvester SCN by 2 and 3, and then divided using two reconfigurable, high-impedance fractional voltage dividers, after which the resulting voltage levels are compared to battery SCN voltages to find the optimal harvesting configuration. During the subsequent harvesting time period, the battery and harvester SCNs are connected according to the optimal configuration determined during the monitoring period, and both processor and battery are powered by the harvesting unit. The monitoring/harvesting cycle is repeated every 6 seconds to allow the PMU to adapt to changing harvesting conditions. If there is insufficient harvested energy, the PMU automatically disconnects the harvester SCN from the battery SCN. The voltage dividers can be configured by the processor to adjust to harvesting sources with different power profiles. There is a trade-off between loss in harvesting time and response time to harvesting condition change with the monitoring/harvesting cycle time.

A. DC-DC Down-Conversion Operation

During down conversion, the PMU converts the battery voltage to two voltage domains($V_{DD1} = 0.6$ Vand $V_{DD2} = 1.2$ V, nominally) to power the processor and an array of peripherals. The proposed PMU has two operating modes with vastly different power budgets: 1) sleep mode when the processor is inactive has a current draw of 1–10 nA and 2) active mode when the processor is running has a current draw of 1–10 μA. To efficiently perform DC-DC down-conversion, the PMU uses four different oscillators: a start-up sequence clock (5 kHz, simulated), sleep mode clock (340 Hz, measured), active mode clock (335 kHz, measured), and an additional SCN clock (3.125 kHz, measured) for harvesting in sleep mode. To reduce ripple magnitude on the output voltage levels at a given SCN frequency and capacitance, a two-way phase interleaved ladder is used as shown in Fig. 8.

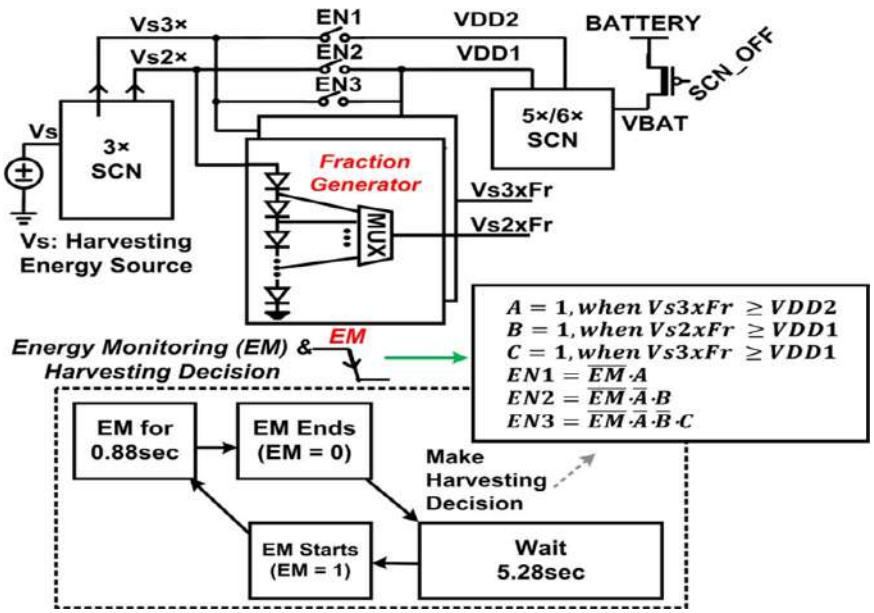
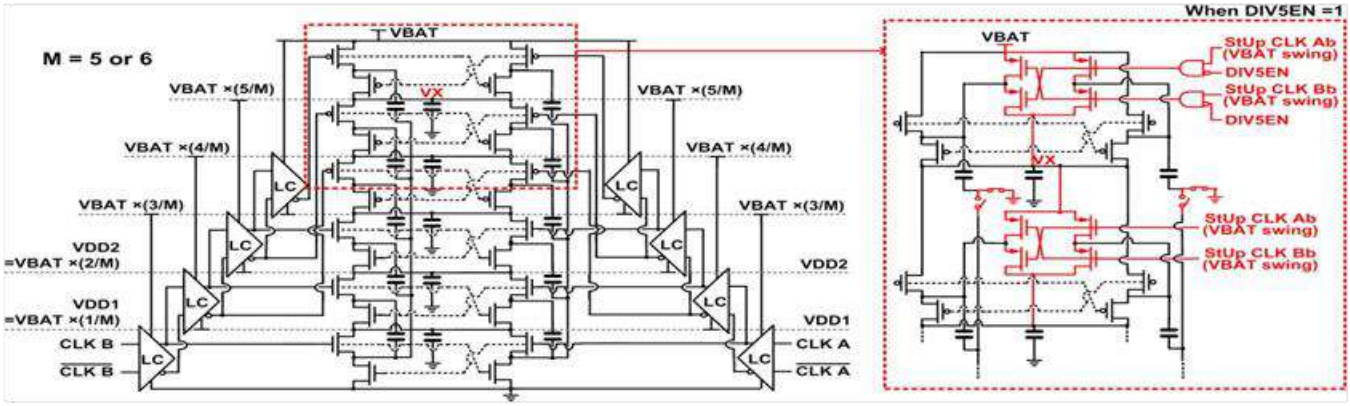
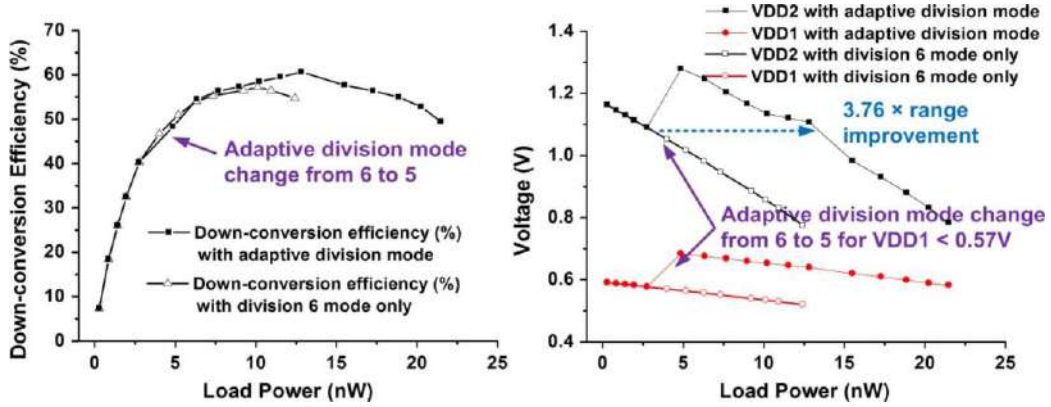


Fig. 7. Proposed overall PMU diagram and adaptive harvesting technique.



(5×/6×)

Fig.8 Two-wayphase –interleaved ladder- type SCN with variable conversionratio.Thick-gateoxideswitchesforstart-upsequenceareindicatedusing arrow FET notation.



Measureddown-conversionef ficiencyinsleepmodewith $\mathrm{VBAT} = 3.6$ Vand340Hzsleepmodeclock(left).MeasuredVDD2andVDD1withadaptive divisionmode,andwithdivision6modeonly(right).

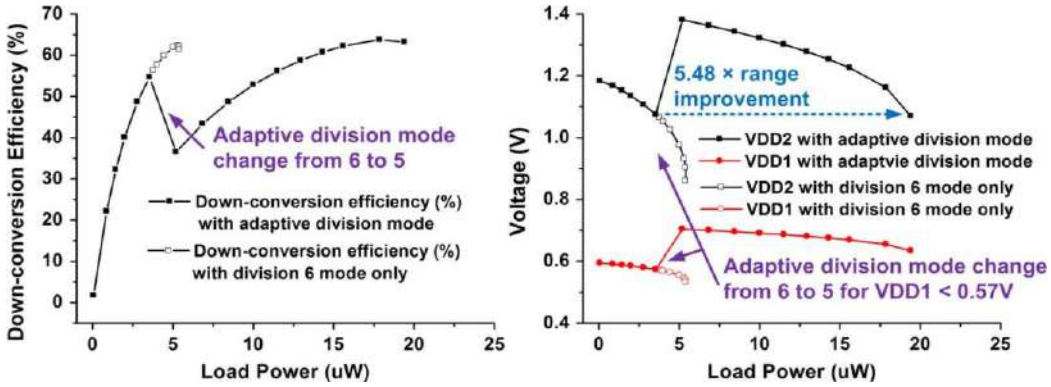


Fig.10. Measureddown-conversionef ficiencyinactivemodewith $\mathrm{VBAT} = 3.6$ Vand335kHzactivemodeclock(left).MeasuredVDD2andVDD1with

The PMU switches between active and sleep mode based on control signals from the processor and timers. The battery has an adaptive conversion ratio of 6 or 5 to address battery voltage variation and a wide range of load power. When 5× is enabled, switches reconfigure the top stage of the ladder such that VBAT and VX are shorted and the bottom plate of the flying capacitor is grounded. This reconfiguration is triggered by two comparators that compare fractions of VDD1 to a voltage [16]. Hysteresis is introduced between the transition to prevent oscillation between the 6 and 5 modes.

The measured power consumption of the sensor platform ranges from 3 nW in sleep mode and up to 20 W in active mode, which can vary with system configuration. Figs. 9/10 shows the down conversion efficiency of the proposed PMU which is nearly equal in sleep mode compared to active mode, due to the use of the VDD1 voltage domain for the SCN clocks in sleep mode. Figs. 9/10 also shows how the configurability of the battery SCN increases the load power range by 3.76 5.48 for sleep/active modes over the case without this capability (VBAT 3.6 V). The battery SCN conversion ratio changes from 6 to 5 when VDD1 drops below 0.57 V, and reverts back to 6 when VDD1 exceeds 0.71 V.

B. PMU Up-Conversion for Adaptive Harvesting

During harvesting, the harvester SCN up-converts the harvesting source voltage by 2 and 3 , and connects either of these outputs to the VDD1 or VDD2 ports of the battery SCN (Fig. 7) The two SCNs are connected in three different ways to accommodate harvesting sources with output voltages from 360–800 mV. The open-circuit voltage develops during the monitoring phase, which lasts 0.88 sec. During this phase a highly resistive diode-connected fraction generator (with configurable output voltage from 0.2 × to 0.95×) is applied to both the 2 and 3 outputs of the harvester SCN. These two factional OCVs are then compared to VDD1 and VDD2 from the battery SCN according to three possible configurations (Confi#1: , Config #2: 2× = VDD1, Config #3: The three configurations are rank ordered according to the harvesting voltage that they induce and a configuration in which the fractional voltage minimally exceeds the voltage from the battery SCN is automatically selected to ensure near-optimal energy

extraction from the harvesting source.

The proposed method automatically adjusts the harvesting setting in response to the harvester's V_{OC} (addressing different harvesting conditions), the battery voltage (addressing different battery charge states), and the setting of the fractional generator (addressing different harvesting sources). For example, when battery voltage is 3.0 V and the OCV fraction is 0.65, configuration 1 corresponds to an of 307–462 mV, configuration 2 to 462–615 mV, and configuration 3 to 615 mV. In a situation with mV, the harvester is automatically disabled to prevent possible drain through the harvesting unit. Finally, after the monitoring phase, the chosen configuration is established during the harvesting phase, which lasts 5.28 s.

$$V_{OC}$$

$$V_{OC} < 307$$

The proposed approach was tested with three harvesting sources: solar cell, microbial fuel cell (MFC), and thermoelectric generator (TEG). The proposed PMU achieves an overall energy harvesting efficiency of 39.8% with an MFC, and 28.1% with a 2.56 cm TEG and 26.9% with a 1.62 mm solar cell. To demonstrate the proposed method for maximum power point tracking (MPPT), we show in Fig. 11(a) how the harvest voltage forced by the PMU tracks the optimum harvesting voltage. As light intensity increases from office light to outdoor lighting, the PMU automatically switches from the first configuration (with force harvest voltage of 200 mV) to the second configuration (300 mV). Fig. 11(b) compares the power extracted from the solar cell by the PMU with the maximum attainable power and the extraction efficiency

$$\frac{\text{Extracted power from solar cell by PMU}}{\text{Attainable power at Maximum Power Point}}.$$

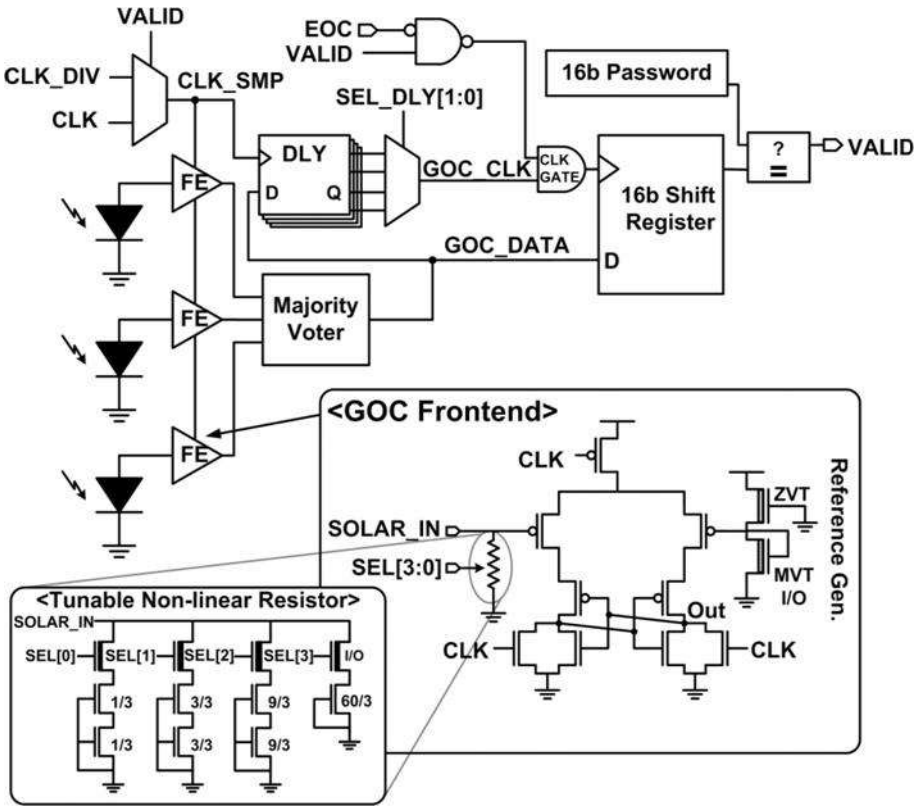


Fig. 11 – GOC Circuit Diagram

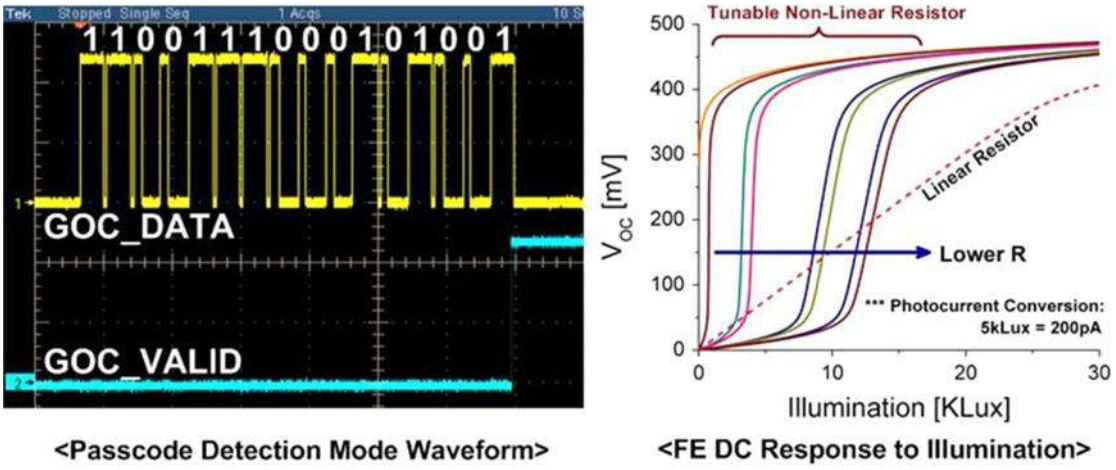


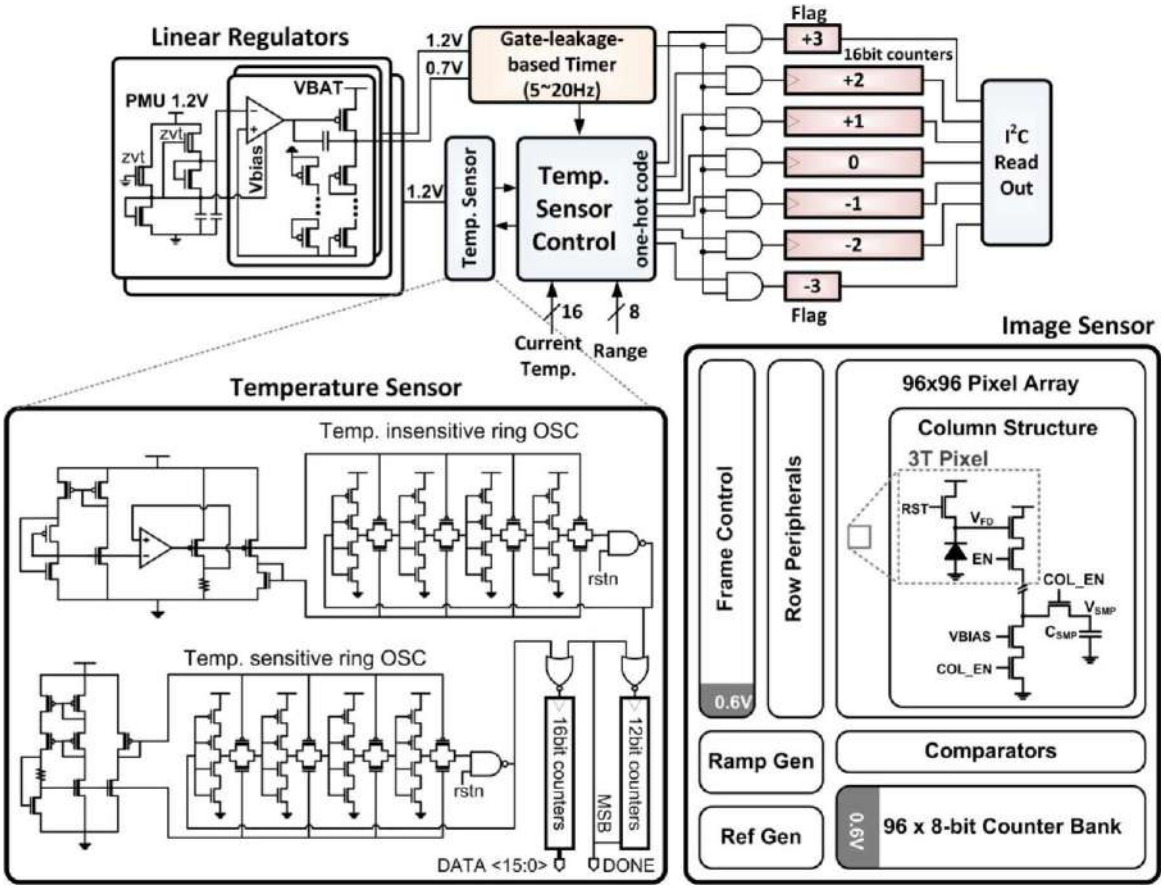
Fig. 12 – Measure waveform of GOC Operation

For light intensities from office light to full daylight, the average extraction efficiency is 93.5%, with a maximum efficiency of 99.2% and minimum efficiency of 76.1%. When the solar cell is exposed to direct sunlight, Voc increases to 520 mV and the PMU switches to the third configuration (400 mV). However, the current at this light condition reaches up to 100 μA (>2 orders of magnitude higher than in office lighting), exceeding the conversion capability of the PMU. Hence the PMU can no longer force the intended 400 mV harvest voltage, resulting in a drop in extraction efficiency.

C. Battery Voltage Monitoring

A battery voltage monitoring unit [14] is included to detect changes in battery voltage and to allow the PMU to take appropriate action. Fig. 12 shows PMU state transitions in response to battery voltage fluctuations. Once the battery voltage stabilizes above 3.4 V, the PMU enters *Operational* mode and the system is activated. Discharging the battery below 3.0 V due to heavy use or wear-out can incur permanent damage to the Li battery.

To prevent this, the monitoring unit detects voltage levels below 3.1 V and the system enters a 185 pW *Deep Sleep* mode where all supplies are turned off. The availability of harvested energy is monitored in Deep Sleep mode and when sufficient light is detected, the system enters *Recovery* mode, recharging the battery from the harvested energy. After the battery has reached a sufficiently high voltage (>3.4 V), the system returns to its normal Operational mode and distributes a power on reset (POR) signal to all IC layers by sequentially releasing the 1.2 V and 0.6 V supplies. This sequence is detected by the POR circuit in each IC layer, which forces local reset for proper initialization.



V. GLOBAL OPTICAL COMMUNICATION

GOC serves three critical purposes that enhance the usability of this sensing platform: initial programming after system assembly, re-synchronization during use, and re-programming out of Deep Sleep mode or when the program has become corrupted. At 228 pW standby power, GOC consumes 20,000 less power than typical RF wakeup radios (4.4 W in [15]), which rely on an un-corrupted software stack (defeating the purpose of re-programming). The GOC module consists of a main control block and three redundant front-end receiver circuits for robustness (Fig. 13). The front-end consists of a photodiode, a pull-down resistor for faster response time, and a comparator. The tunable resistor is intentionally implemented with off-state MOSFETs, whose non-linear resistance profile improves light detection sensitivity by 14× compared to a linear resistor, as seen in Fig. 14. The comparator compares the photodiode's anode voltage to a reference voltage of approximately 200 mV, generated by a 2-T reference generator [16], to determine whether the light level corresponds to a "1" or a "0". The outputs of the three comparators are majority-voted, and this digitized signal, GOC_DATA, is then delayed by a tunable delay chain to generate GOC_CLK. Subsequently, GOC_CLK's rising edge samples GOC_DATA to determine data "1" or "0".

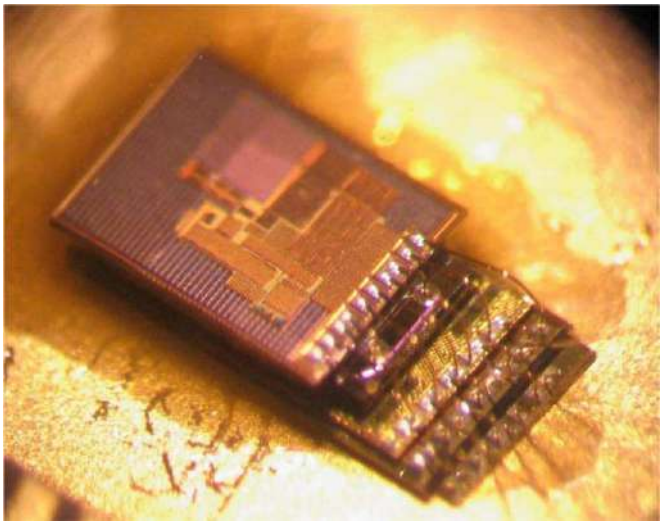
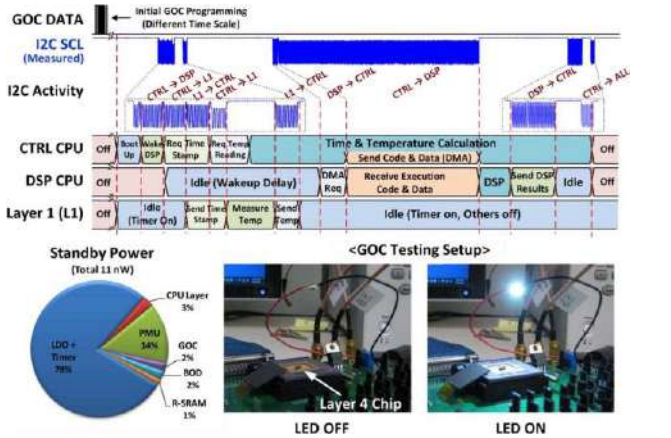


Fig. 14 Photo of proposed 1.0 mm Sensing Platform

Since the GOC is located in Layer 4, the front-end receivers are placed between bonding pads for light exposure. To prevent false triggers from ambient light, a 16-bit predetermined bit pattern is used as a global passcode to initiate a GOC transaction. Once the passcode is validated, GOC runs at an 8× faster rate to enable higher transmission rate while avoiding the high standby mode power associated with such a clock. Additionally, a local chip-ID/masking scheme allows for selective batch-programming of different groups of sensor nodes. GOC is measured to be operational up to 120 bps and consumes 72 pJ/bit.



Measured wave form of the I^2C SCL linealongwiththecorrespondingusagescenario (top).Measuredstandbypowerdistribution (bottomleft)and Testings etupfor GOC(bottomright).

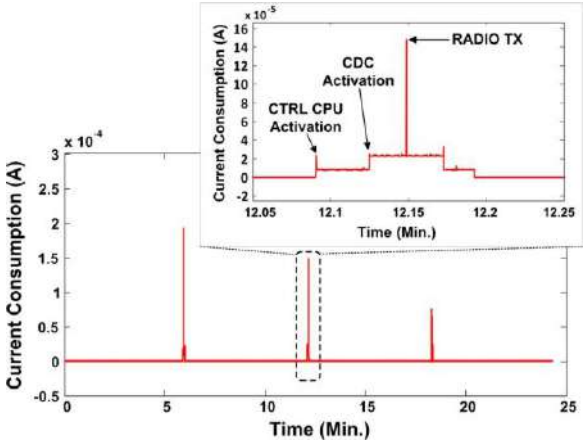


Fig. 15 Measured current waveform of a periodically operating sensor node with a radio layer.

VI. TIMER AND SENSORS

A gate leakage based timer and temperature monitor provide temperature-compensated time tracking, which is critical for synchronizing wireless communication. The timer [10] is augmented with a temperature sensor (Fig. 15) so that temperature compensation and/or calibration can be provided. The temperature sensor includes a reference current generator and a PTAT (proportional to absolute temperature) current generator. The reference current generator provides a constant current over temperature while the PTAT current generator gives current proportional to temperature. Voltage from those current generators is fed to two separate ring oscillators in order to convert current to frequency. The ring oscillator consists of a chain of stacked inverters with transmission gates between the inverters. Voltage from the current generator drives the gates of the transmission gates, changing their resistance and setting the ring oscillator frequency. Since the driving current controlled by the reference current generator does not change over temperature, the frequency of its oscillator is also constant. The output of the temperature insensitive ring oscillator is counted by a 12-bit counter, and its MSB is used to measure the fixed time. For this constant time, a 16-bit counter counts the output of the sensitive one. Thus, temperature information is stored in the 16-bit counter in digital format. This temperature sensor is measured to consume 806 nJ/pt with a standard deviation of 0.51 C.

Also shown in Fig. 15 is a low power, dual-supply 96 96 imager that uses $\mathrm{psub/n+}$ parasitic diodes in the standard CMOS process as photodiodes. A conventional 3-T pixel structure is used to achieve a high pixel fill-factor of 74%. The column source followers are biased in sub threshold to limit the tail currents, and a pulsed control signal activates the source followers only for a short period. The pixel array and the comparators in the 8-bit single-slope ADC use 1.2 V supply for better signal-to-noise ratio, whereas the digital components use 0.6 V supply to minimize switching power. The image sensor consumes 680 nJ/frame.

VII. SYSTEM OPERATION

We demonstrate and provide measurement results of the reconfigurable platform implemented in two different systems. The first system is the baseline system consisting of three die whose micrograph is shown in Fig. 16: 1) imager, timer, solar cells; 2) control processor, PMU and 3) DSP processor. Measured waveforms of the SCL wire in I^2C in Fig. 17 shows the communication activity among the layers and the state transition graph at bottom describes the operation of each layer. In this trace, dies were connected using board-level connections that are identical to the connections in the stacked system. Also the dies themselves were identical to those shown in the stacked system of Fig. 2, with the exception of added observability circuits to aid measurement. After initial shipping from manufacturing, the sensor node is in sleep mode and can be programmed through GOC as shown by the 'GOC DATA' waveform in Fig. 14. Control CPU Layer then initiates a boot up sequence and also wakes up the DSP CPU Layer. As shown, timestamp and temperature measurement request is sent to Layer 1 and results are transferred back to the Control CPU Layer for temperature-calibrated timestamp calculation. Meanwhile, DSP execution code and data is transferred to the DSP CPU Layer while the Control CPU Layer is performing its operation. This way, the Control CPU Layer and the DSP CPU Layer can concurrently process data. When the DSP Layer completes, its result is sent back to the Control CPU Layer to be stored in retentive memory located on the Control CPU Layer. The Control CPU then places the entire system into standby mode so that the full system consumes minimum power until the next active operation for periodic sensor measurement.

Fig. 17 also shows the power budget of the system in standby mode. Total standby power consumption is11 nW with the dominant portion consumed by the gate-leakage-based timer. The timer is the only nW-level active unit in the entire system in standby mode and is required to provide an accurate timing reference. Without any harvesting, the integrated 0.6 μAh thin-film battery can support the system in sleep mode for up to 2.3 days. For applications where accurate timing is not required, standby power can be reduced to 2.4 nW with a pW timer [17], allowing 10.5 days of sleep mode operation without energy harvesting.

By taking advantage of the re-configurability of the proposed sensor system, we also demonstrate a second system where the DSP processor layer is swapped with a new layer that includes a capacitance-to-digital converter (CDC) and a near-field radio [8], while all other layers are reused. This configuration represents a sensor node that periodically wakes up to take pressure sensor measurements and transmits measured data through a radio for monitoring of pressure in tumors as an early diagnostic tool for determining chemo-therapy effectiveness Operation of the second system is measured as shown in Fig. 18. The sensor node wakes up every 6 minutes and has current consumption on the order of 3 nA in standby, 5 A for active mode without radio transmission, and peaks to 10 A when radio transmission is active. Note that the differences in the current peak height are due to the limited sampling frequency in the long-term trace, which leads to some samples being reported higher than others.

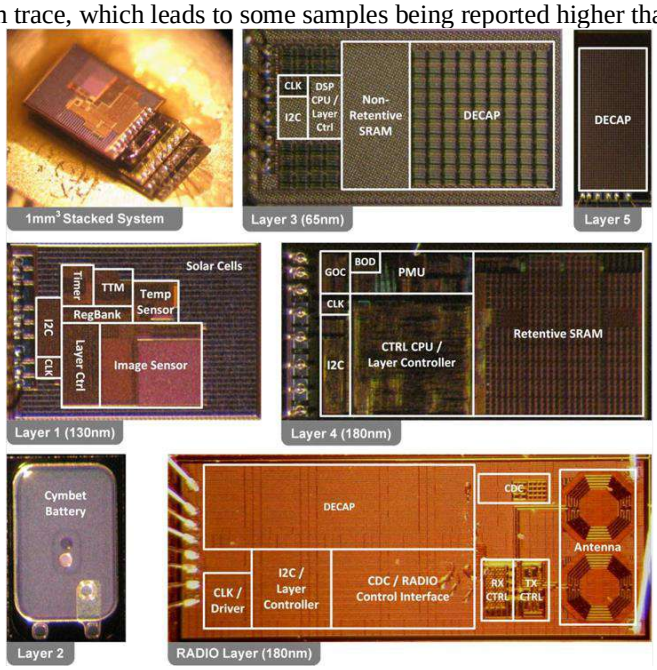


Fig. 16 shows the micrograph of the proposed 1.0 mm sensor system and chip micrograph of each layer in the system.

VIII. CONCLUSION

After reviewing the papers on mm-scale sensors we came to know that power plays an important role for the implementation of the smart dust technologies like mm-scale sensors which are of the size of dust grain but still have the computational power. Through this review paper we tried to show that how can we change the circuits in 1.0mm3 sensor for reducing the heat emission in complex computational operations and harvesting the energy so that the life of the sensor extends. We did this by using modular die-stacked structure that allowed easy extension of mm-scale sensors. The multi-modal energy harvesting scheme allows to harvest energy from wide range of energy resources and using a batter-voltage dependent power management scheme that allowed safe operation performing without excessive battery-discharge. I2C was used in the physical deployment so that we could establish proper communication between the layers and also make this 1.0mm3 sensor compatible for further extensions. A low power optical communication scheme allows energy efficient programming and synchronization of sensor nodes. These circuit techniques or circuit structure create an ultra low power sensor platform, having two microprocessors, 19kB memory; and low power image sensors and timers in 1.0 mm3 volume. We can use this circuit diagram for future implementation of energy efficient mm-scale sensors application and can finally advance our vision of technology to the point where smart dust technologies can be implemented and we can integrate further technologies in it, specially making it intelligent through ML and AI.

REFERENCES

[1] G. Bell, "Bell's Law for the birth and death of computer classes," *Commun. ACM*, vol. 51, no. 1, pp. 86–94, Jan. 2008.

[2] T. Nakagawa, G. Ono, R. Fujiwara, T. Norimatsu, T. Terada, M. Miyazaki, K. Suzuki, K. Yano, Y. Ogata, A. Maeki, S. Kobayashi, N. Koshizuka, and K. Sakamura, "1-cc computer: Cross-layer integration with UWB-IR communication and locationing," *IEEE J. Solid-State Circuits*, vol. 43, no. 4, pp. 964–973, Apr. 2008.

[3] Y. Tachwali, H. Refai, and J. Fagan, "Minimizing HVAC energy consumption using a wireless sensor network," in *33rd Annual Conf. IEEE Industrial Electronics Society*, Nov. 2007, pp. 439–444.

[4] N. Mohamed and I. Jawhar, "A fault tolerant wired/wireless sensor network architecture for monitoring pipeline infrastructures," in *Int. Conf. Sensor Technologies and Applications*, Aug. 2008, pp. 179–184.

[5] N. Elvin, N. Lajnef, and A. Elvin, "Feasibility of structural monitoring with vibration powered sensors," *Smart Materials and Structures*, vol. 15, pp. 977–986, June 2006.

[6] L. Schwiebert, S. Gupta, and J. Weinmann, "Research challenges in wireless networks of biomedical sensors," in *Int. Conf. Mobile Computing and Networking*, 2001, pp. 151–165.

[7] B. Warneke, M. Last, B. Liebowitz, and K. S. J. Pister, "Smart dust: Communicating with a cubic-millimetercomputer,"*Computer*, vol.34, pp. 44–51, Jan. 2001.

[8] G. Chen, H. Ghaed, R. Haque, M. Wieckowski, Y. Kim, G. Kim, D. Fick, D. Kim, M. Seok, K. Wise, D. Blaauw, and D. Sylvester, "A cubic-millimeter energy-autonomous wireless intraocular pressure monitor," in *IEEE ISSCC Dig. Tech. Papers*, 2011, pp. 310–311.

[9] E. Y. Chow, S. Chakraborty, W. J. Chappell, and P. P. Irazoqui, "Mixed-signal integrated circuits for self-contained sub-cubic millimeter biomedical implants," in *IEEE ISSCC Dig. Tech. Papers*, 2010, pp. 236–237.

[10] Y. Lee, B. Giridhar, Z. Foo, D. Sylvester, and D. Blaauw, "A 660 pW multi-stage temperature-compensated timer for ultra-low-power wireless sensor node synchronization," in *IEEE ISSCC Dig. Tech. Papers*, 2011, pp. 46–47.

[11] G. Chen, M. Fojtik, D. Kim, D. Fick, J. Park, M. Seok, M. Chen, Z. Foo, D. Sylvester, and D. Blaauw, "Millimeter-scale nearly perpetual sensor system with stacked battery and solar cells," in *IEEE ISSCC Dig. Tech. Papers*, 2010, pp. 288–289.

[12] L. Chang, Y. Nakamura, R. K. Montoye, J. Sawada, A. K. Martin, K. Kinoshita, F. H. Gebara, K. B. Agarwal, D. J. Acharyya, W. Haensch, K. Hosokawa, and D. Jamsek, "A 5.3 GHz 8 T-SRAM with operation down to 0.41 V in 65 nm CMOS," in *IEEE Symp. VLSI Circuits Dig. Tech. Papers*, 2007, pp. 252–253.

[13] NXP Semiconductors, "I^2C-bus specification and user manual, UM10204 datasheet, Rev. 3," Jun. 2007.

[14] I. Lee, S. Bang, Y. Lee, Y. Kim, G. Kim, D. Sylvester, and D. Blaauw, "A 635 pW battery voltage supervisory circuit for miniature sensor nodes," in *IEEE Symp. VLSI Circuits Dig. Tech. Papers*, 2012, pp. 202–203.

[15] K. Yadav, I. Kymissis, and P. R. Kinget, "A 4.4 μW wake-up receiver using ultrasound data communications," in *IEEE Symp. VLSI Circuits Dig. Tech. Papers*, 2011, pp. 212–213.

[16] M.Seok,G.Kim,D.Sylvester,and D.Blaauw,"A 0.5 V 2.2pW 2-transistor voltage reference," in *Proc. IEEE CICC*, 2009, pp. 577–580.

[17] Y.-S. Lin, D. Sylvester, and D. Blaauw, "A sub-pW timer using gate leakage for ultra low power sub-Hz monitoring system," in *Proc. IEEE CICC*, 2007, pp. 397–400.

ADGITM, New Delhi AT A GLANCE

Dr. Akhilesh Das Gupta Institute of Technology & Management (ADGITM), formerly Northern India Engineering College, a premier academic Institute of Delhi, INDIA is recognized by AICTE and is affiliated to GGS Indraprastha University, Delhi. ADGITM has been established by BBDES, Lucknow, UP in the year 2003 and it offers B.Tech (CSE, IT, ECE, EE, ME & CE) and MBA. Every year, students are procuring Gold Medals in academics as well as in sports and co-curricular activities. The students have brought laurels to the college by winning a total of 25 GGSIPU Gold Medals and continuously striving to bring fame to the institute.

Dr. Akhilesh Das Gupta Institute of Technology & Management
(Formerly Northern India Engineering College)
FC-26, Shastri Park, New Delhi
Ph. 011-49905900-99
Website: www.adgitmdelhi.ac.in

ISBN: 978-93-5493-753-8

978-93-5493-753-8