David C. Wyld,
Jan Zizka (Eds)

Computer Science & Information Technology

6th International Conference on Signal and Image Processing (SIGI 2020),
December 26 - 27, 2020, Chennai, India.

AIRCC Publishing Corporation

Volume Editors

David C. Wyld,
Southeastern Louisiana University, USA
E-mail: David.Wyld@selu.edu

Jan Zizka,
Mendel University in Brno, Czech Republic
E-mail: zizka.jan@gmail.com

ISSN: 2231 - 5403
ISBN: 978-1-925953-34-3
DOI: 10.5121/csit.2020.102001- 10.5121/csit.2020.102008

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

The 6th International Conference on Signal and Image Processing (SIGI 2020), December 26 - 27, 2020, Chennai, India, 6th International Conference on Computer Science and Information Technology (CSTY 2020) and 6th International Conference on Artificial Intelligence and Applications (AI 2020) was collocated with 6th International Conference on Signal and Image Processing (SIGI 2020). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from theWest.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The SIGI 2020, CSTY 2020 and AI 2020 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was doneelectronically.

In closing, SIGI 2020, CSTY 2020 and AI 2020 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the SIGI 2020, CSTY 2020 and AI 2020.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students andeducators continues beyond the event and that the friendships and collaborations forged willlinger and prosper for many years tocome.

David C. Wyld,
Jan Zizka (Eds)

General Chair	Organization
David C. Wyld,	Jackson State University, USA
Jan Zizka (Eds)	Mendel University in Brno, Czech Republic

Program Committee Members

Abdallah Makhoul,	University Bourgogne - Franche-Comte, France
Abdel-Badeeh M. Salem,	Ain Shams University, Cairo, Egypt
Abhijit Das,	Manipal University, India
Abhishek Shukla,	A P J Abdul Kalam Technical University, India
Abid Ismail,	University of Kashmir, India
AdeyanjuSosimi,	University of Lagos, Nigeria
Adnan A. Rawashdeh,	Yarmouk University, Jordan
Ahmad Al-Aiad,Jordan	University of Science and Technology, Jordan
Ahmed A. Elngar,	Beni-Suef University, Egypt
Ahmed M. Soliman,	Cairo University, Egypt
Akhil Gupta,	Lovely Professional University, India
Alejandro Regalado Mendez,	Universidad del Mar, Mexico
Alessandro Massaro,	Dyrecta Lab, Italy
Ali Al-Sabbagh,	ministry of communication, Iraq
Alia Karim Abdulhassan,	University of Technology, Iraq
Alvaro Rocha,	University of Coimbra, Portugal
AnasAlSobeh,	Yarmouk University, Jordan
Antonio Abreu,	ISEL, Portugal
AssemA.Hamied Moussa,	SRC Africa, Egypt
Atanu Nag,	Modern Institute of Engineering & Technology, India
Ayman A. Aly El-Naggar,	TaifUniversity, Saudi Arabia
Ayman Kassem,	Cairo University, Egypt
Azeddine WAHBI,	Hassan II University, Morocco
B. F. Oladejo,	University of Ibadan, Nigeria
Baihua Li,	Loughborough University, UK
BarhoumiWalid,	SIIVA-LIMTIC Laboratory, Tunisia
Basanta Joshi,	Tribhuvan University, Nepal
Bilal H. Abed-alguni,	Yarmouk University, Jordan
Bin Zhao,	Xidian University, China
BoukariNassim,	skikda university, Algeria
BRAHAMI Menaouer,	National Polytechnic School of Oran, Algeria
CallistusIreneousNakpih,	St. John Bosco's College of Education, Ghana
Changsoo Je,	Sogang University, South Korea
Chien-Cheng Yu,	Hsiuping University of Science and Technology, Taiwan
Dac-Nhuong Le,	Haiphong University, Vietnam
DadmehrRahbari,	University of Qom, Iran
Deepak Mishra,	ISRO, India
Deepali Gupta,	Maharishi Markandeshwar University, India
Dharmendra Sharma,	University of Canberra, Australia
DiabAbuaiadah,	Waikato institute of technology, New Zealand
DibyaMukhopadhyay,	University of Alabama, US
Dinesh Bhatia,	North Eastern Hill University, India
Dinesh Kumar Saini,	Sohar University, Oman
Dinh-Thuan Do,	Eastern International University, Vietnam

K.Lenin,	Prasad V. Potluri Siddhartha Institute of Technology
Kabid Hassan Shibly,	Dhaka International University, Bangladesh
Karim MANSOUR,	University Salah Boubenider, Algeria
Ke-Lin Du,	Concordia University, Canada
keyvanpour,	Alzahra University, Iran
khaoulaboutouhami,	southeast university, Nanjing china
Khin Su Myat Moe,	Yangon Technological University, Myanmar
Kin-Choong Yow,	University of Regina, Canada
Klimis Ntalianis,	Athens University of Applied Sciences, Greece
Koh You Beng,	University of Malaya, Malaysia
Korchiyne Redouan,	Ibn Tofail University, Morocco
Lucian Lupu-Dima,	University of Petrosani, Romania
M. Kamala Kumari,	AdikaviNannaya University, India
Manal Mostafa Ali,	Al-Azhar University, Cairo
Mansour Y. Bader,	Al-Balqa Applied University, Jordan
Marzan Tasnim Oyshi,	Daffodil International University, Bangladesh
Maumita Bhattacharya,	Charles Sturt University, Australia
Meera Ramadas,	Amity University, India
Mihai Lungu,	University of Craiova, Romania
Mohamed Ali El-sayed,	Benha University, Egypt
Mohamed ArezkiMellal,	University of Maryland, USA
Mohamed Fakir,	Sultan MoulaySlimane University, Morocco
Mohamed Issa,	Zagazig University, Egypt
Mohamed SahbiBellamine,	Carthage University, Tunisia
Mohammad Abido,	King Fahd University, Saudi Arabia
Mohammad Ashraf Ottom,	Yarmouk University, Jordon
Mohammed Benyettou,	University center of Relizane, Algeria
Mohammed GH. AL Zamil,	Yarmouk University, Jordan
Mohd.Rizwan beg,	R B Group of Institutions,Agra, India
Mueen Uddin,	Effat University Jeddah, Saudi Arabia
Muhammad Arif,	Guangzhou University, China
Muhammad Sajjadur Rahim,	University of Rajshahi, Bangladesh
Muhammad Sarfraz,	Kuwait University, Kuwait
Murat Karabatak,	Firat University- Elazig/Turkey
Mu-Song Chen,	Da-Yeh University, Taiwan
Mustafa Abuzaraida,	Universiti Utara Malaysia, Malaysia
Nawaf Omar N. Alsrehin,	Yarmouk University, Jordan
Nawres KHLIFA,	University of Tunis El Manar, Tunisia
Nguyen Dinh Lau,	University of Danang, Vietnam
Nidal Turab,	Al-ahliyya Amman University, Jordan
Niloofar Rastin,	Shiraz University, Iran
Nouh Elmitwally,	Jouf University, KSA
Olakanmi Oladayo O,	University of Ibadan, Nigeria
Oleksii K. Tyshchenko,	University of Ostrava, Czech Republic
Parameshachari B D,	Jain University, India
Pi-Chung Hsu,	Shu-Te University, Taiwan
Ramadan Elaiess,	University of Benghazi, Libya
Ramakrishnan,	DR.MGR Educational And Research Institute, India
Ramgopal Kashyap,	Amity University Chhattisgarh, India
Ravinder Reddi,	Chaitanya Bharathi Institute of Technology, India
Ray-I Chang,	National Taiwan University, Taiwan
Reguia Mahmoudi,	Yahia Fares University of Medea, Algeria

Technically Sponsored by

Computer Science & Information Technology Community (CSITC)

Artificial Intelligence Community (AIC)

Soft Computing Community (SCC)

Digital Signal & Image Processing Community (DSIPC)

Organized By

Academy & Industry Research Collaboration Center (AIRCC)

TABLE OF CONTENTS

6th International Conference on Signal and Image Processing (SIGI 2020)

6th International Conference on Computer Science and Information Technology (CSTY 2020)

6th International Conference on Artificial Intelligence and Applications (AI 2020)

IMAGE PROCESSING FAILURE AND DEEP LEARNING SUCCESS IN LAWN MEASUREMENT

J. Wilkins[1], M. V. Nguyen[1] and B. Rahmani[1, 2]

[1]Math and Computer Science Dept., Fontbonne University, St. Louis, MO, USA
[2]Computer Science and Data Science Dept.,
Maryville University, St. Louis, MO, USA

ABSTRACT

Lawn area measurement is an application of image processing and deep learning. Researchers used hierarchical networks, segmented images, and other methods to measure the lawn area. Methods' effectiveness and accuracy varies. In this project, image processing and deep learning methods were used to find the best way to measure the lawn area. Three image processing methods using OpenCV compared to convolutional neural network, which is one of the most famous, and effective deep learning methods. We used Keras and TensorFlow to estimate the lawn area. Convolutional neural network or shortly CNN shows very high accuracy (94-97%). In image processing methods, thresholding with 80-87% accuracy and edge detection are the most effective methods to measure the lawn area while the method ofcontouring with 26-31% accuracy does not calculate the lawn area successfully. We may conclude that deep learning methods, especially CNN, could be the best detective method comparing to image processing learning techniques.

KEYWORDS

Lawn Measurement, Convolutional Neural Network, Thresholding, Edge Detection, Contouring.

1. INTRODUCTION

While there are definitely major business applications for lawn care companies, this project is a great experiment into an area of image processing that is rarely looked at: image regression. So, the project had to be started nearly from scratch, and creative ways were found to deal with the limited quality and quantity of the samples.

At the lawn care company, the lawn area of houses from satellite images had to be measured. The process is often rather simple: find the address of the house, use online software to measure the area of the lawn of the house, record the measurements, and repeat the steps. While the company sends out people to manually measure the lawn area, they want some rough estimate beforehand in areas. It is a very repetitive job, but it could only be done by a human. So, the lawn care companies have to pay the person at least minimum wage, and the person has to spend many hours trying to measure hundreds of houses. Therefore, the company is spending thousands of dollars and inordinate amount of time to measure houses. Eye-strain and fatigue are another issue that causes inconsistency and low accuracy. The lawn care company could save thousands of dollars and hundreds of human hours in measuring the lawns of houses accurately (at least within a margin of error).

David C. Wyld et al. (Eds): SIGI, CSTY, AI - 2020
pp. 01-10, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.102001

Weiqi Zhou and his colleagues in 2008 tried to measure the lawn properties including lawn and house area remotely. They used hierarchical networks and classified segmented images for this purpose [1]. Alexander Schepelmann in his master thesis, measured the lawn area using color and visual texture classifier [2]. One of the methods that could be used by the lawn care company is artificial neural networks. Artificial neural networks have a long winding history in the computer science field. While the idea of a neural network in general has its basis in biology and did influence artificial neural networks, their mathematical conception dates back to 1943 in a paper of Warren S. McCullough and Walter Pitt [3]. In 1950s, several attempts were made to simulate a neural network by Clabaugh, et. al. In 1959, Stanford researchers Bernard Widrow and Marcian Hoff successfully implemented it in an algorithm [4]. They removed echoes and other data corruption from phone lines by predicting the next actual bit. In the 1980s, neural networks were revived as a topic of interest, and several advances to the artificial neural network architecture were made. After that productive decade, progress slowed down; however, due to even further developments in both artificial neural networks and increased processing power in computers, neural networks are now extremely popular and making large advances (Hardesty). Relevant to our query about measuring lawns from satellite images, convolutional neural networks are a major advancement in recent years for a variety of reasons. Specifically, convolutional neural networks would work theoretically well because the lawn measurement problem is an image problem. The convolutional neural network pulls out the features such as grass texture, tree patterns, and other image features, the dense network layers analyze those features and determine the lawn area.

In this project, first we show the disability of image processing methods in calculating house and lawn area, then show that deep learning and convolutional neural network would work very well in this purpose. Data will be described in section 2. Section 3 discusses the theoretical background. Results and conclusions come after in section 4.

2. DATA DESCRIPTION

Creating the neural network for the deep learning problem is only half of the battle. The other half is collecting all the necessary data. The specific dataset of satellite images of houses with their measured lawn area is a big challenge. While there were datasets that involved satellite imagery, they are often involved much larger areas and usually implemented in classification problems. Despite this problem, there have been imperfect but still effective ways of collecting the data needed for this project. The most immediate solution is measuring the lawn area of houses on Google Maps and cropping these images out of the screenshots.

While image quality was somewhat lower, it was the best dataset that could be found. We used an online software Area Calculator (Area Calculator Using Maps) to measure each total lawn area in square meters. Then we used Krita, an image editing software (“Digital Painting. Creative Freedom”) to crop the larger picture and make smaller individual pictures. The resulted images contain one of the measured areas with its surrounding lawn. In total, the author collected 65 pictures of houses and their lawns. This dataset was too small to actually use as training data. In order to remedy this problem, we used the ImageDataGenerator class mentioned below in order to duplicate each picture. In many deep learning image processing problems, artificial data is a legitimate way to reduce potential overfitting and increase performance. One way is to duplicate a picture but then slightly change it in some way in order to produce an essentially different picture; this can be done by image’s rotating, inverting or flipping, distorting, changing the brightness of the image, and more. The ImageDataGenerator class has a method that can do all of those manipulations by just specifying in the constructor parameters of how you want to change each picture. The change is often random and within a given parameter range. This class was used to iterate through each picture in the dataset and create an augmented duplicate of each

image 50 times. Figure 1 is a sample of duplicating images. 3000 images created with this method. We took a portion of created dataset as a test dataset. Test data for validation determines the overfitting.

3. THEORETICAL BACKGROUNDS

3.1. Convolutional Neural Network

The convolutional neural network (CNN) helps to achieve a basis in deep learning projects. It is popular for classification rather than regression that is the main part of the current problems. Most importantly, CNN was included in the Keras library, which has been used to carry out a great deal of the process. Ultimately, the neural network did not reach the goal of 90% accuracy, but the accuracy was between 85% and 90%. Unfortunately, it did not necessarily explore overfitting in-depth and still did not apply to the current work [5, 6]. Hence, we used Tensorflow that was more complex but too low-level to easily implement.

Figure 1. Left-Image of before Data Augmentation; Right-Image of after Data Augmentation

There are many ways to implement convolutional neural networks. Sometimes, it works by defining and implementing complex functions with code from scratch [7] or by using Keras with Tensorflow as a backend [8]. A convolutional neural network is a 2D network meaning that it could take in 2-dimensional data. This fact is the kind of neural network used for image processing because pictures represented as 2-dimensional through their height and width.

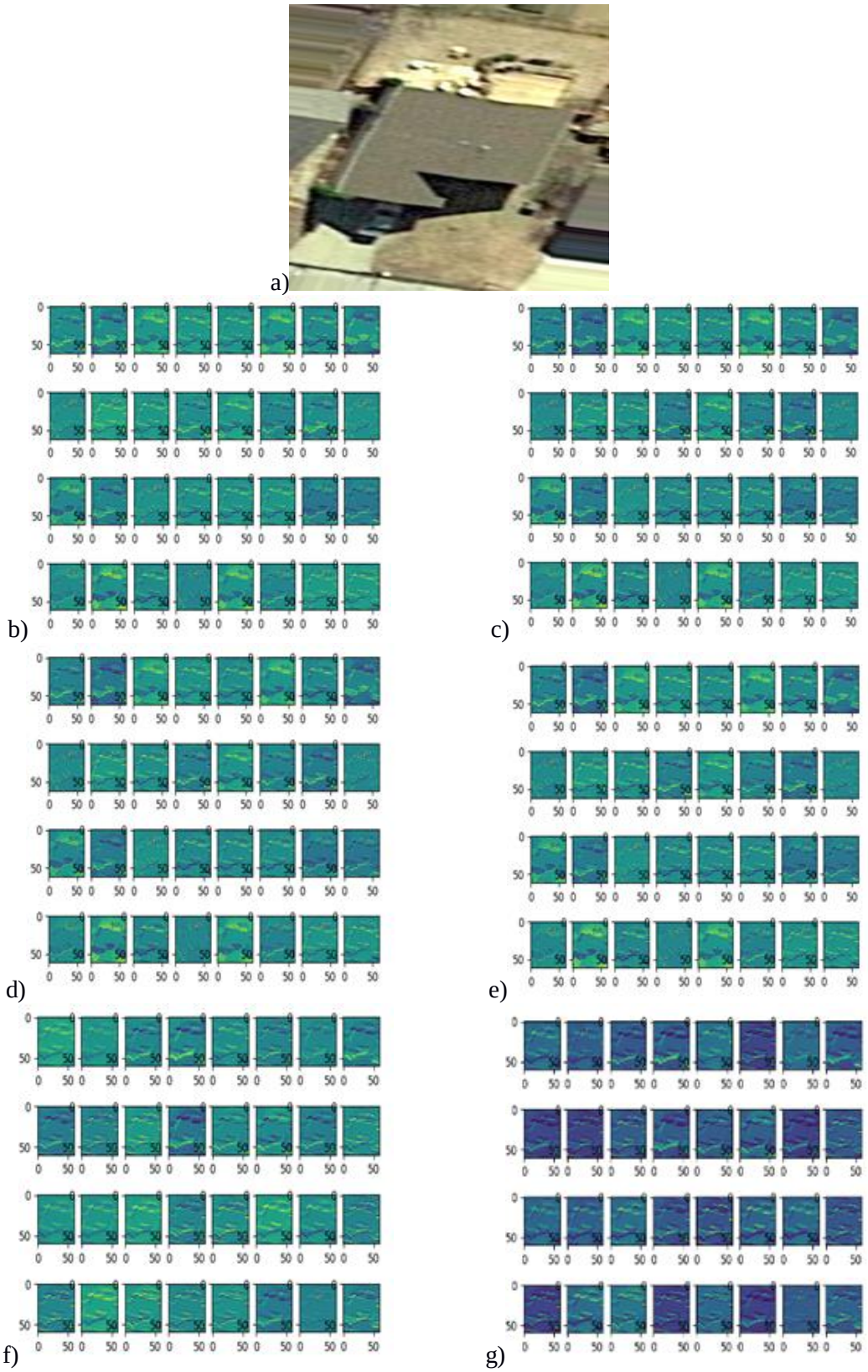

Figure 2. a) The original picture; b) Activation layer 1; c) Activation layer 2; d) Activation layer 3; e) Activation layer 4; f) Activation layer 5; g) Activation layer 6

However, the inputted arrays are 4 dimensions because it also includes the number of color channels and the number of pictures. It is 1-dimensional CNN for sentences and 3-dimensional CNN for voices. We started with three to four convolutional layers with two dense output layers at the bottom to process the information from the previous layers. After a few iterations, we used a different version of the convolutional layer in Keras that could be supposedly better performance called the separable convolutional layer. The separable convolutional layer or depth wise separable convolutional layer is an advanced version of the convolutional layer that can often increase the performance of the deep learning image processing models [3].

Besides, we implemented a new activation function in each layer called an ELU (instead of the previous RELU function) which again is often associated with producing better performance in neural networks [2]. Then, we implemented Batch Normalization to increase performance significantly, reduce the overfitting in the model, and deactivate a given layer in a certain percentage of the time [2]. We apply regularizations at each layer to prevent overfitting [9]. We also added the usual convolutional layer complements such as pooling layers to reduce the size of each picture at each layer [2]. The final model after intensive training on several models was a nine-layer convolutional neural network with 6 convolutional and 3 dense layers. This idea is probably the only neural network that specialized in regression analysis with the given satellite dataset. The basic procedure is as following: An array containing all the picture's arrays passed through the first convolutional layer, which pulls out certain low-level features from the picture. Then, the ELU activation function determines what values/features should move to the next convolutional layer. A pooling layer reduces the size of the picture and passes from every two convolutional layers. After passing through the six convolutional layers, the resulted features as a flattened list, processed by the three dense layers, which determine the area, correlated with the given feature [10]. By pulling out the features of the pictures through these convolutional layers and activations, the model simplified the data and emphasized certain features. Figure 2 shows an example for results of this process. Besides, the fine-tuning of the neural network, which involves trial and error attempts of a different number of neurons, layers, regularize' parameters, and optimizers, was also included in this example.

In fine-tuning of the neural network, many of parameters were tested with a class from the Scikit-Learn library in Python called GridSearchCV (Sklearn.model_selection.GridSearchCV) [11]. GridSearchCV is a way of automating the trial and error process as well as testing the true accuracy or performance of a neural network. One of the concepts in this library is cross-validation or CV. Essentially, CV takes the existing dataset and separates it into further sub-sets that can be used for validation. The GridSearch part of GridSearchCV tests data by iterating through hyper-parameters. Hence, the GridSearchCV finds the best parameter too. While this method served its purpose well, it limited the dependability; specifically, in reducing overfitting. Because the validation partly relied on duplicated data from the same dataset, it does not show the best accuracy for reducing overfitting. Still, it was useful to see which parameters would over-fit the most and then try to reduce it by adjusting the dropout functions which calculates manually. Overall, it led to the best performance as of now. Figure 3 shows the final cropped model.

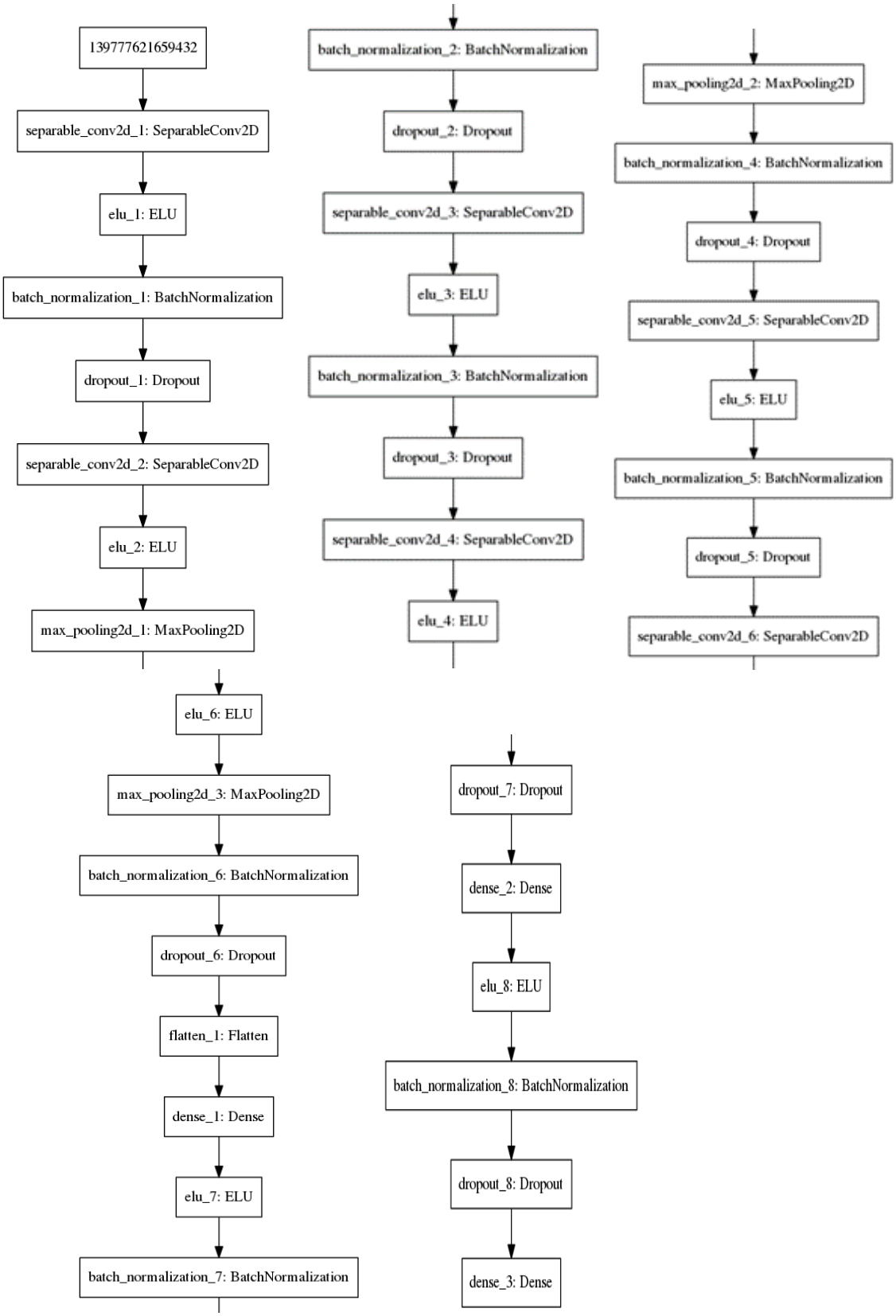

Figure 3. The final cropped images of the main model used, the model starts with the most left image and ends at the most right image

3.2. Thresholding and Augmentation

Though the model significantly fine-tuned, it was still below 90% accuracy or even got to 85% especially on the test set. Therefore, we worked on another solution like OpenCV, which is a non-deep learning solution to the problem [12]. OpenCV is an image-processing library in Python. It has several methods and classes used to manipulate images that can extract desired features. OpenCV emphasizes or removes unnecessary features from the pictures, which lets the neural net focuses on the essential features for measuring the lawn area. The 3 different studied methods include thresholding, contouring finding, and edge detection [13]. Thresholding requires a range of pixel color densities and converts them to black and white images. Its purpose is to make the house in the middle of each picture completely change to one color (white) and the rest of the lawn area to another color (black). Threshold eliminates functions unrelated to lawn area, thus simplifying the measurement of lawn area. Figure 4 shows an image before, after Thresholding, and after both Thresholding and Augmentation.

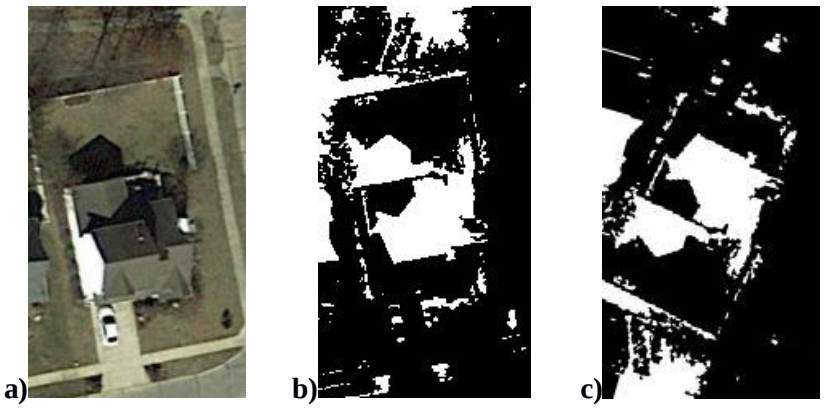

Figure 4. a) Image before Thresholding; b) Image after Thresholding;
c) Image after Thresholding and Augmentation

3.3. Contouring and Augmentation

After several tests and some fine-tuning, Thresholding and Augmentation did not work very well because they could not even replicate the current model's accuracy. The next method was finding the contours of the houses and their lawn area. The contouring method finds the edges and curves of the houses (Figure 5). Despite adding extra information contouring also failed to exceed the performance of the original neural network.

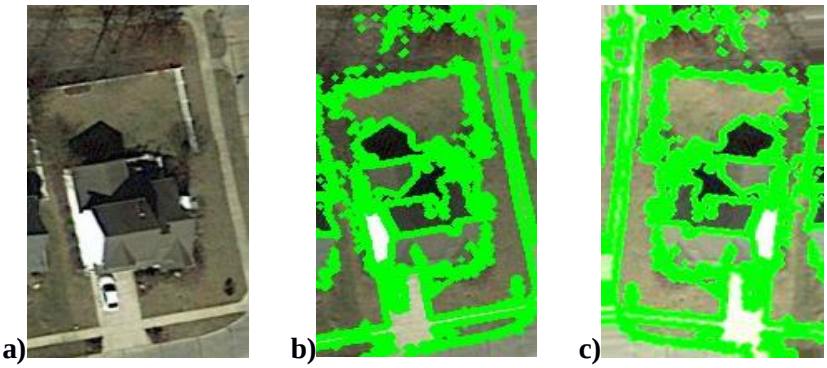

Figure 5.a) Image before Contouring; b) Image after Contouring and Gaussian Blurring; c) Image after Contouring and Augmentation

3.4. Canny Edge Algorithm

Finally, we used the Canny Edge algorithm to find the edges of the house traced by the program that was able to simulate the desired result of the thresholding (Figure 6). Unfortunately, this algorithm also failed to meet expectations. Despite OpenCV being a powerful tool in many image processing applications, it was insufficient for this project's purpose.

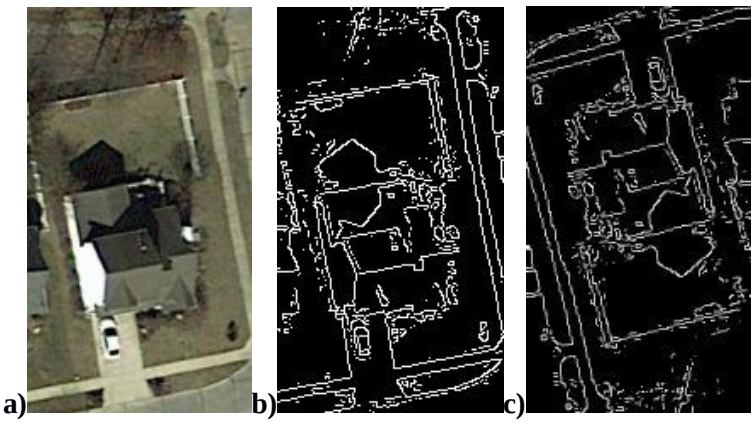

Figure 6. a) Image before Canny Edge Algorithm; b) Image after Canny Edge Algorithm; c) Image after Canny Edge Algorithm and Augmentation

4. RESULTS AND CONCLUSION

Table 1 shows the accuracy of different models. The final accuracy of convolutional neural network is not quite at 90%, but it definitely was able to measure lawn area within a given margin of error. Looking at the Mean Squared Error (MSE) of the prediction of the test data after the most recent training on 1,849 pictures, MSE of 1437 is the given result which is around 38 square meters margin of error for each predicted lawn area ($\sqrt{1437}$ = ~38 sq. m.). To get something like an accuracy percentage, the ratio of the margin of error to the actual average value subtracted from 1 (1 - (error/average value)). The average measurement of the data that was trained on this model (pictures 1-45) is roughly 294 sq. m. The accuracy is on average ~87% (1-(38/294) = ~87%). The median, which is approximately 276 sq. m. with the accuracy: ~86% (1-(38/276) = ~86%). The test result, unfortunately, is the result of some overfitting that has been occurring, but it is not that radically different from the training results. The mean squared error of the training vs the testing is only off by a less than 100 MSE even though there is a greater difference between training and validation MSE.

However, there is some evidence of volatility in this model; for example, in another test, the MSE was 2366 meaning an average error of 49 Squared meters, ~83% average accuracy, and ~83% median accuracy. While this would not be acceptable to any lawn care business yet, this is still a sign of hope that this neural network can at least predict the value within a reasonable margin of error.

As mentioned before, there are 65 pictures accumulated. So, there are approximately 3000 duplicated pictures in total. However, the final result used only the 45 original pictures which were separated into training, validation, and test data. In this case, there were really 1,849 training samples, 150 validation samples, and 250 test samples when this neural network achieved the above accuracy. The results obtained by image processing methods have almost half the accuracy of deep learning methods. Our results could be a resource to help people working on

properties lawn area and other business companies [14,15,16].

Table 1. Mean Squared Error of training, validation, and testing datasets.

Model Used	**Highest Accuracy (1-Error/Average of Original Data)**	**Model Results (Average Predicted Lawn Area)**	**Average Lawn Area of Used Data**
CNN Training	-	-*	~298.97 m^2
CNN Validation	~94%	~280.12 m^2	~298.14 m^2
CNN Testing	~97%	~262.65 m^2	~254.17 m^2
Threshold Model Training	-	-*	~283.55 m^2
Threshold Model Validation	~80%	~228.87 m^2	~287.21 m^2
Threshold Model Testing	~87%	~221.22 m^2	~254.17 m^2
Contour Model Training	-	-*	~297.66 m^2
Contour Model Validation	~26%	~634.49 m^2	~280.35 m^2
Contour Model-Testing	~31%	~587.50 m^2	~254.17 m^2
Edges Model Training	-	-*	~297.89 m^2
Edges Model Validation	~85%	~237.80 m^2	~280.53 m^2
Edges Model Testing	~57%	~145.16 m^2	~254.17 m^2

∗ The average predicted result was not shown for the training data since the MSE recorded during training does not correspond with the MSE and the predicted lawn area observed through Keras predict function for the training data

REFERENCES

[1] W. Zhou, A. Troy, M. Grove (2008) Modeling Residential Lawn Fertilization Practices: Integrating High Resolution Remote sensing with Socioeconomic Data.Springer.

[2] A. Schepelmann (2010) Identification and Segmentation of Lawn Grass Based on Color and Visual Texture Classifiers, Master Thesis, Case Western Reserve University.

[3] W.S. Mcculloch, W. Pitts (1990) A Logical Calculus of the Ideas Immanent in Nervous Activity, Bulletin of Mathematical Biology Vol. 52, No. l/2. pp. 99-115. 1990.

[4] V. Reddy, M. K. Kumar, K. A. Naik, Widrow Hoff (2011) Learning Algorithm Based Minimization of BER Multi-User Detection in SDMA, International Journal of Technology And Engineering System (IJTES): Jan –March 2011- Vol. 2 No. 3.

[5] Geron (2017) Hands-on Machine Learning with Scikit-Learn and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems, O'Reilly.

[6] F. Chollet (2017) Deep Learning with Python, Manning.

[7] Vedaldi, K. Lenc (2016) MatConvNet Convolutional Neural Networks for MATLAB, ArXiv: 1412.4564v3.

[8] R. Airola, K. Hager (2017) Image Classification, Deep Learning and Convolutional Neural Networks, Karlstad University.

[9] S. Demyanov (2015) Regularization Methods for Neural Networks and Related Models, The University of Melbourne.

[10] M. Heusel, D.A. Clevert, G. Klambauer, A. Mayr, K. Schwarzbauer, T. Unterthiner, S. Hochreiter(2015) ELU-Networks: Fast and Accurate CNN Learning on ImageNet, ICCV 2015.

[11] Scikit-learn developers (2018) Scikit-learn User Guide, Release 0.19.2 June 2018.

[12] A. Mordvintsev, K. Abid (2017) OpenCV-Python Tutorials Documentation, Release 1, 2017.

[13] R. C. Gonzalez, R. E. Woods (2017) Digital Image Processing, Pearson.

[14] Project Administration Handbook for Civil Engineering Works, (2018).

[15] E. Schappacher (2018) Embracing web-based tools to measure lawns, Landscape Management.

[16] P. Robbins, T. Birkenholtz (2002) Turfgrass revolution: measuring the expansion of the American lawn, Land Use Policy 20 (2003) 181–194, Pergamon.

BIOMETRIC FOETAL CONTOUR EXTRACTION USING HYBRID LEVEL SET

Rachana Jaiswal and Srikant Satarkar

Department of Computer Science and Engineering,
Amravati University, Akola, MH, India

ABSTRACT

In medical imaging, accurate anatomical structure extraction is important for diagnosis and therapeutic interventional planning. So, for easier, quicker and accurate diagnosis of medical images, image processing technologies may be employed in analysis and feature extraction of medical images. In this paper, some modifications to level set algorithm are made and modified algorithm is used for extracting contour of foetal objects in an image. The proposed approach is applied on foetal ultrasound images. In traditional approach, foetal parameters are extracted manually from ultrasound images. Due to lack of consistency and accuracy of manual measurements, an automatic technique is highly desirable to obtain foetal biometric measurements. This proposed approach is based on global & local region information for foetal contour extraction from ultrasonic images. The primary goal of this research is to provide a new methodology to aid the analysis and feature extraction from foetal images.

KEYWORDS

Active contour, Region-based, Edge-based, Hybrid Approach, Foetal images.

1. INTRODUCTION

Image segmentation plays significant role in medical image processing and computer vision. It permits visualization of organs and structures in the body, analysis of the shapes and sizes of these organs to clinicians, and diagnosis of pathologies more quickly and accurately. Physical segmentation is a tiresome, time consuming process and most of the times it is not accurate. It is subjected to high variations from one expert to another due to its dependence on physician's subjective knowledge and his experience. The information generated from the computerized analysis of medical images can be used as a second opinion in making diagnostic decisions by physicians.

Foetal biometric ultrasound measurements pose an interesting problem due to presence of inhomogeneous intensity profile, strong speckle noise and shadows making it difficult to properly segment the correct positions and shapes of interesting regions. In addition to this, edges are poorly defined and are close to other foetal structures making segmentation task difficult using standard active contour models. The proposed work aims at developing a novel segmentation technique to extract foetal contour accurately with reduced dependence on initial curve placement [9].

David C. Wyld et al. (Eds): SIGI, CSTY, AI - 2020
pp. 11-20, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.102002

2. IMAGE SEGMENTATION TECHNIQUE

There are two approaches to segmentation similarity-based and discontinuity-based. In similarity-based approach, image is partitioned based on uniformity in intensity and texture. These methods create more coherent regions compared to discontinuity-based methods. In this approach, segmentation is performed based on the grey value of pixels without taking account of connectivity property. Thresholding, Region growing methods fall into this category. Discontinuity-based approach partitions image based on sudden intensity changes. At the point of intersection of two regions, intensities vary. In images, object boundaries represent edges. The traditional methods employ Sobel, Robert and Prewitt operators for edge detection. Another discontinuity-based approach involves use of deformable model.

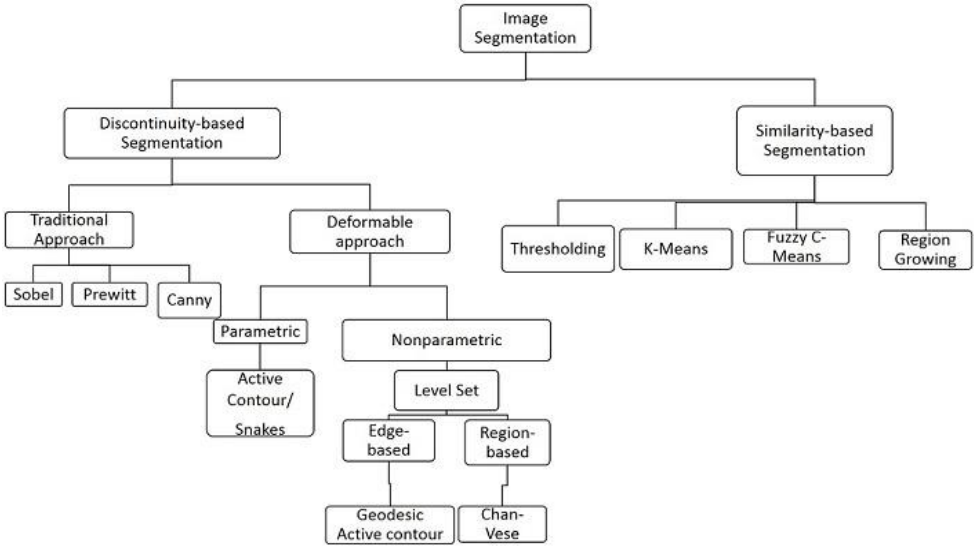

Figure 1: Image Segmentation Algorithm classification

Deformable models possess ability to directly generate closed parametric curves or surfaces from images and are robust to noise and spurious edges [2]. Deformable models can be categorized into the parametric deformable models and the geometric model depending on way of evolution.

Table 1: Comparison of parametric and nonparametric methods

Parametric	Non-parametric/geometric
use explicit representation based on Lagrange formulation	represents the contour implicitly and evolves according to the Euler formulation
Initialization sensitive	Initialization sensitivity is less compared to parametric ones.
Fail to converge to concavities	Affected by weak edges
Top-down approach	Multiple boundaries at a time can be found.
Prior knowledge can be easily incorporated	Are able to handle high curvature regions.
Less human intervention.	reduces the time required for image analysis and subjectivity associated with manual delineations and measurements [6].
Prior boundary information is available.	No prior information is available.

3. IMAGE SEGMENTATION BASED ON LEVEL SET

It is proposed by Stanley Osher and James Sethian. Prior information regarding shape and the initial locations of the region to be segmented is not needed in this approach. It is widely used for medical, satellite and natural image segmentation, analysis and boundary extraction. It can handle corner, inflexion and topological structure changes and can also segment regions with irregular shape. It is initialization insensitive [3]. It can segment objects with deep concavities and can detect multiple objects [4].

Any segmentation problem can be solved by converting it into a PDE framework and employing finite difference methods. Let C represents contour evolving by the time t in the normal vector direction N with the speed function F (K) which depend on the curvature of the contour. The contour deformation using PDE can be written as,

$$\frac{\partial c}{\partial t} = F(K).N$$

In a high curvature area, contour propagation speed is higher compared to a low curvature area [8]. In level set method, speed and direction of the propagation depends on image intensity gradient and the curvature. Internal and external energies along the boundary of curve are minimized by deforming the curve along its shape [11].

For a closed curve C which is dividing plane into two regions, signed distance function (SDF) is given ϕ (x, y). For point lying inside the curve C, ϕ (x, y) > 0; for outside point, ϕ (x, y) < 0; ϕ (x, y) = 0 for the point lying on C [7]. To deal with topological changes naturally following equation is used:

$$\frac{\partial \emptyset}{\partial t} + F(k)|\nabla \emptyset| = 0$$

It is not able to detect objects in images with low-contrast boundaries [5]. The geometric active contour models are classified into edge-based and region-based models. Edge-based active contour models use local edge information for curve evolution whereas region-based active contour models use statistical information. In traditional level set method**,** the leaking problem appears due to presence of a fuzzy or discrete boundary in the region. The iteration time is affected by too large or too small initial contour [13].

Table 2: Edge-based and Region-based Comparison

Edge-based	Region-based
an edge stopping function is based on image gradient.	the motion of the contour relies on region descriptor
Highly sensitive to noise, affected by fuzzy or blurred boundaries.	Insensitive to noise. It is less affected by fuzzy or blurred boundaries.
Affected by initial curve placement. Can segment heterogeneous objects	robustness against initial curve placement. not ideal for segmenting heterogeneous objects.
Uses local edge information	Uses statistical region information
Have high segmentation precision since it allows user to mark boundary directly.	Can detect interior and exterior boundaries simultaneously.

4. LEVEL SET IMAGE SEGMENTATION IMPROVEMENT

Motivated by the work of [17], a hybrid segmentation approach which aims to blend the benefits of global and modified local region-based approaches to improve the quality of segmentation is proposed. It is a two-stage method which uses local and global intensity information.

Region-based level set approach is classified into global and local region-based approach. Initially, contour for foetal image is specified and global region-based approach is applied to it. Global approach is fast and stable with homogeneous regions. But our foetal images contain inhomogeneous regions and contour that is extracted is not proper. Hence, the output of this stage is given as input to modified local region-based approach which employs square window to give exact contour.

MATLAB is used for experimentation on personal computer with Intel Core i5, 1.7 GHz and 4GB RAM. Sufficient medical images taken from authorized sonography centre are tested and after testing, result shows improvement in quality and efficiency of segmentation as compared to existing approaches. The parameters used for comparison of algorithms include Structural Similarity Index, Correlation, Mean Square Error, Peak Signal to Noise Ratio, Dice similarity coefficient, Jaccard Similarity Coefficient. The result of the evaluated values of Structural Similarity Index, Correlation, Mean Square Error, Peak Signal to Noise Ratio, Dice similarity coefficient, Jaccard Similarity Coefficient for some of the images are represented in comparison graph.

Proposed approach when tested over a large dataset of foetal images has given encouraging results. The results which are presented shows Structural Similarity index for different images using proposed approach performs 6.94% better over existing method, for Correlation it is 5.97%, MSE is decreased by 66.06%, Peak Signal to Noise ratio is improved by 16.73%, for Dice Similarity Coefficient it is 5.9% and for Jaccard Similarity Coefficient it is 29.89%. So, it is proposed after careful observations that the hybrid approach gives better performance among all other segmentation techniques.

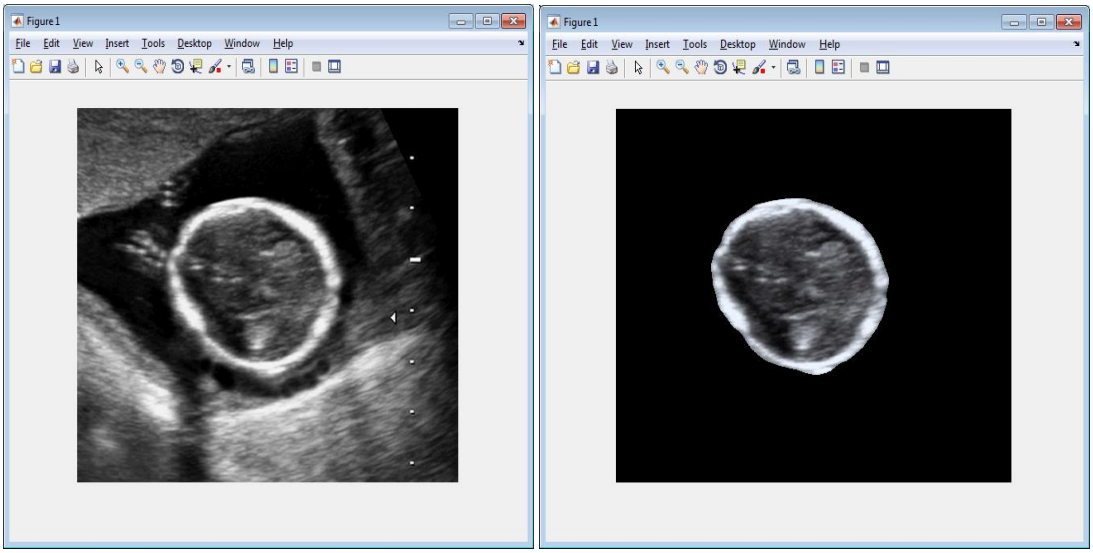

Figure 2: a) I[st] Input Image b) Result of Proposed Hybrid Method

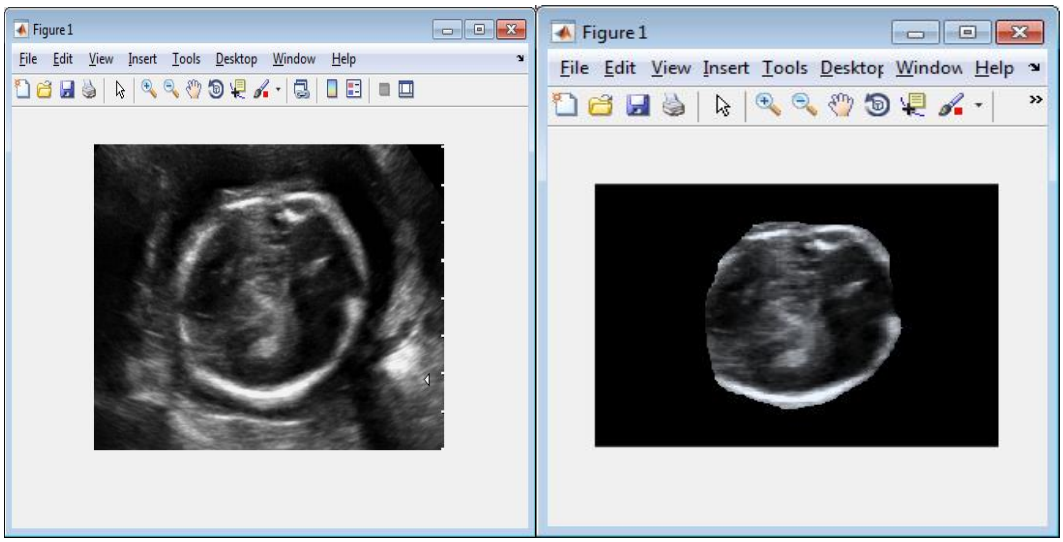

Figure 3: a) II[nd] Input Image b) Result of Proposed Hybrid Method

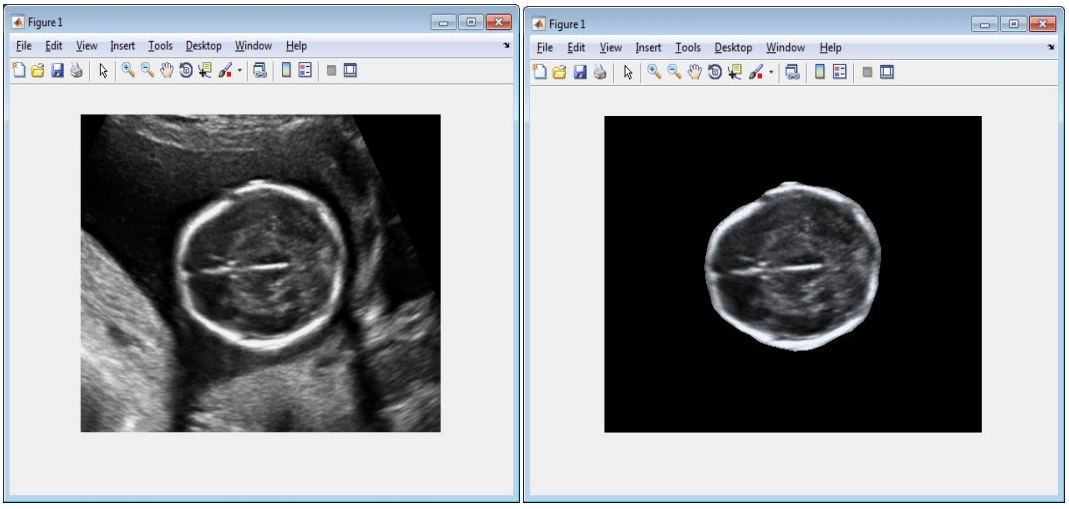

Figure 4: a) III[rd] Input Image b) Result of Proposed Hybrid Method

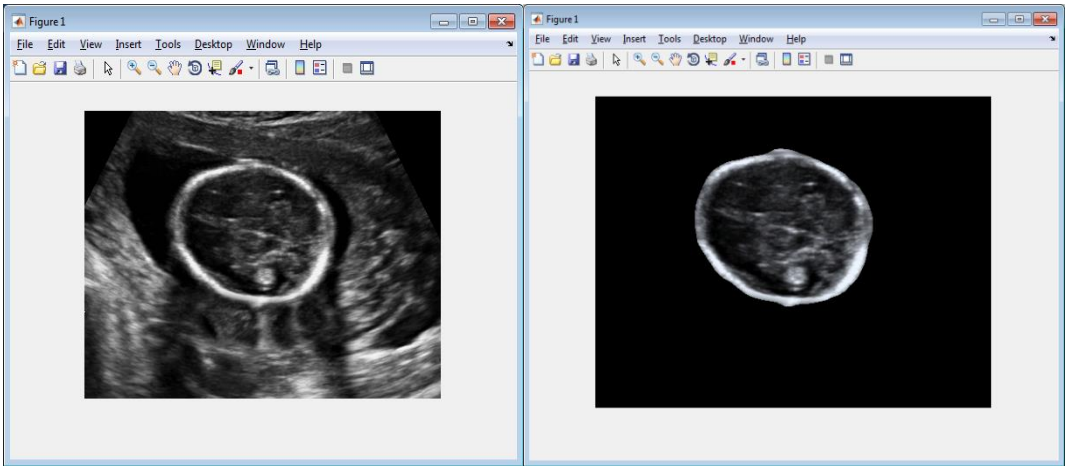

Figure 5: a) IV[th] Input Image b) Result of Proposed Hybrid Method

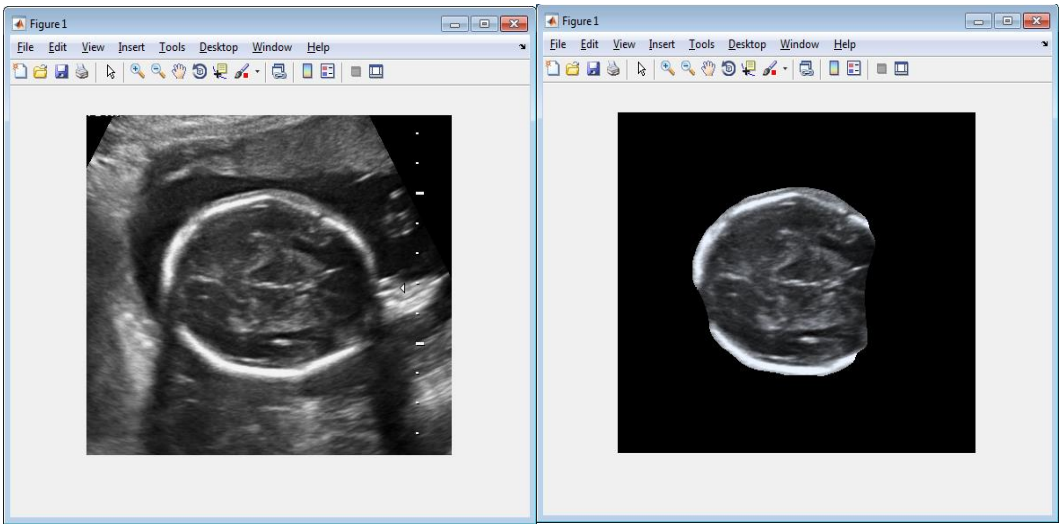

Figure 6: a) V[th] Input Image b) Result of Proposed Hybrid Method

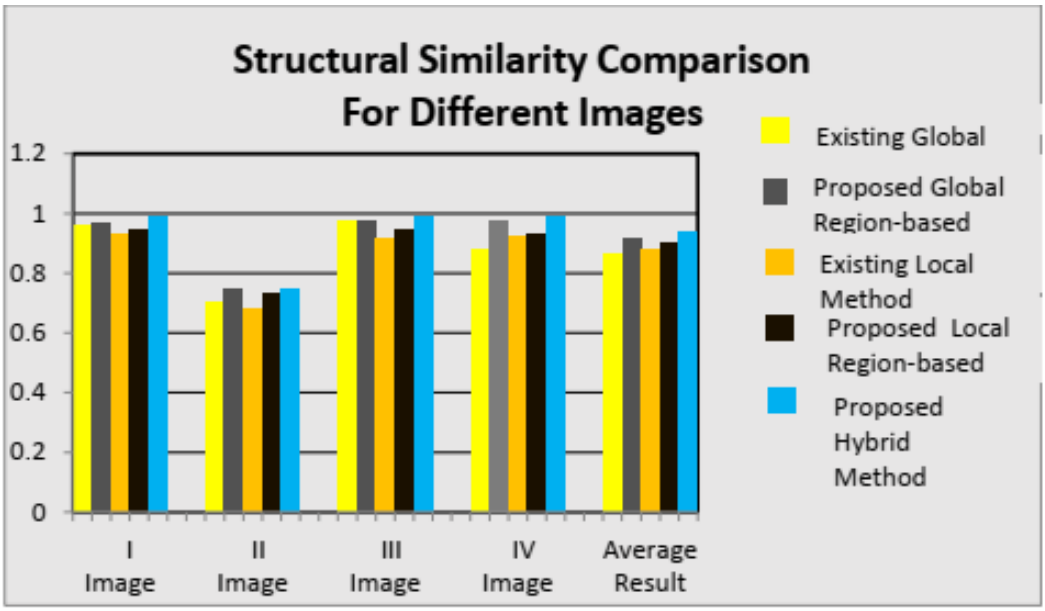

Figure 7: Structural Similarity Comparison for different images

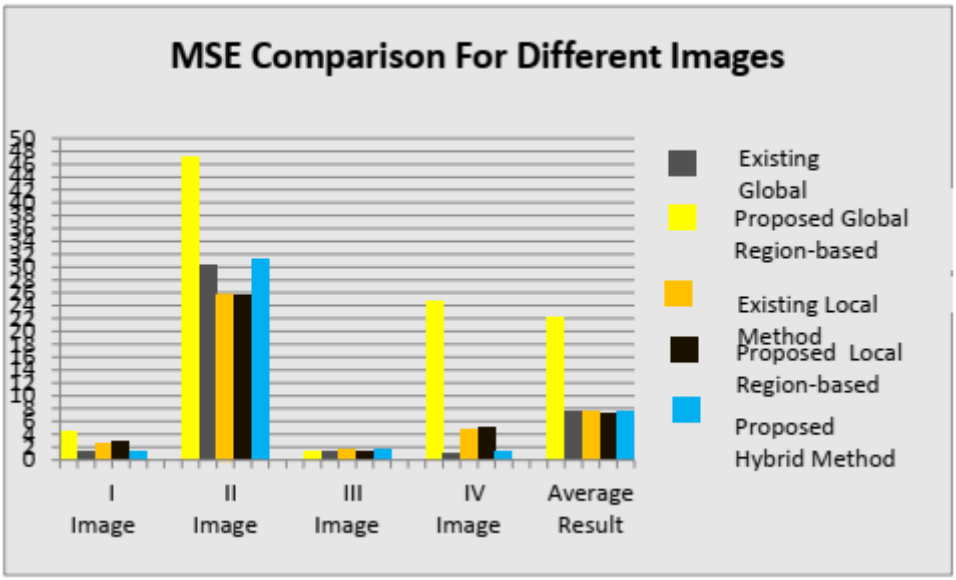

Figure 8: MSE Comparison for different images

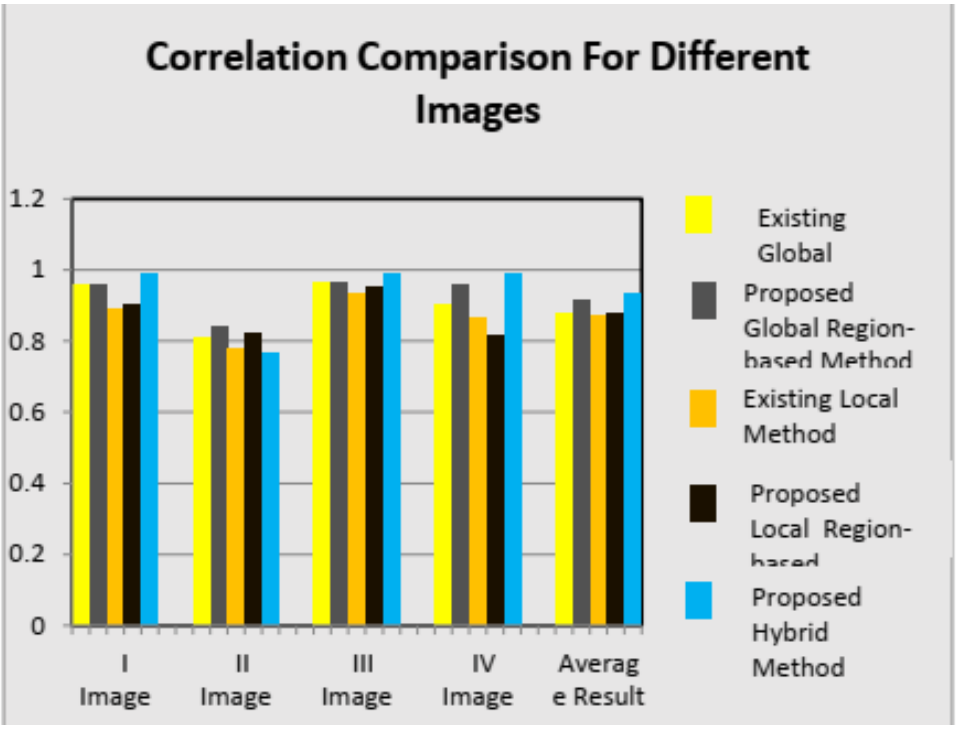

Figure 9: Correlation Comparison for different images

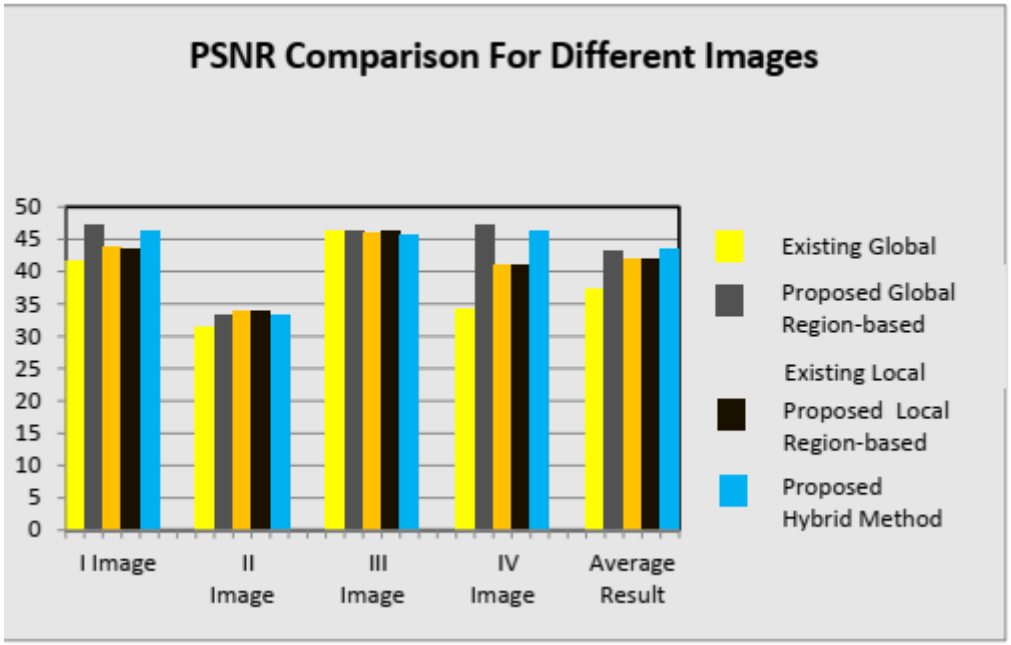

Figure 10: Peak Signal to Noise Ratio Comparison for different images

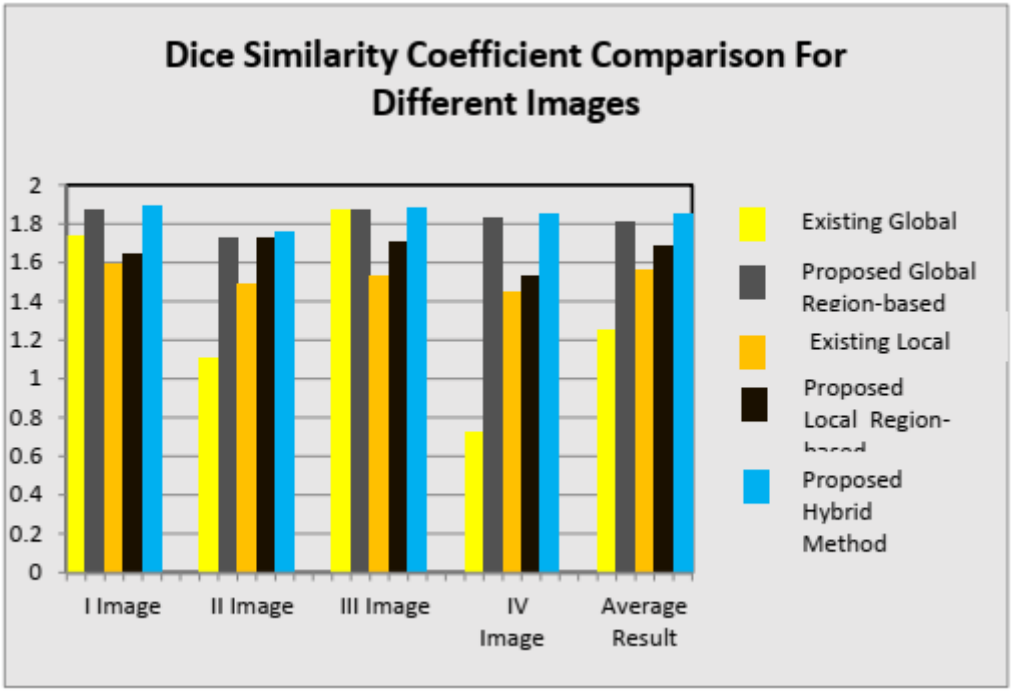

Figure 11: Dice Similarity Coefficient Comparison for different images

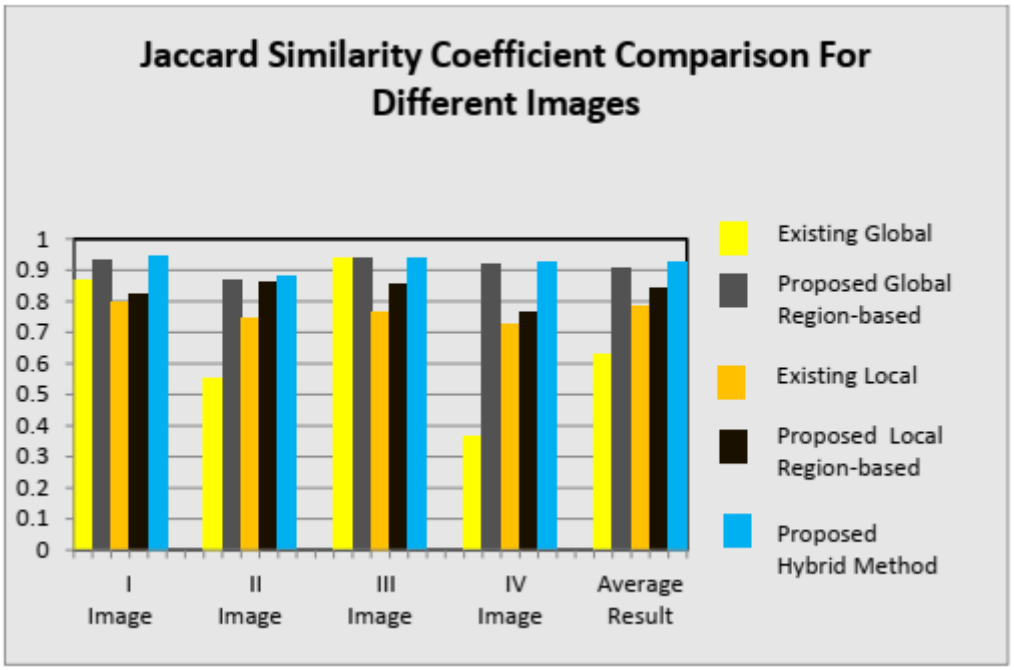

Figure 12: Jaccard Similarity Coefficient Comparison for different images

5. CONCLUSION

Perona and Malik smoothing technique is applied prior to segmentation [18]. The narrow band method is employed in the proposed approach. The proposed global region-based model is insensitive to the location of initial contour. The model could not produce correct result in presence of heterogeneous object. So, after observing the performance of global region-based approach, local region-based approach is implemented. It is a local method, so in presence of heterogeneous object or incomplete edges, this algorithm produces correct result. In order to obtain benefits of both methods, result after application of proposed global method which is rough segmentation is given as input to proposed local method for suitable initialization. The segmentation method is robust to different types of initialization and provides more accurate foetal contour which can be used for further analysis.

REFERENCES

[1] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh and Eero P. Simoncelli, (2004) "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, pp.600–612.

[2] HaiJun Wang, Ming Liu and WenLai Ma, (2010) "Color Image Segmentation Based on a New Geometric Active Contour Model", International Conference on Machine Vision and Human-machine Interface, DOI: 10.1109/MVHI.2010.75, pp. 6 – 9.

[3] Mohamed Ali Charmi, Mohamed Amine Mezghich, Slim M'Hiria, Stephane Derrode and Faouzi Ghorbel, (2010) "Geometric shape prior to region-based active contours usingFourier-based shape alignment", IEEE International Conference on Imaging Systems and Techniques, DOI: 10.1109/IST.2010.5548489, pp.478 – 481, 2010.

[4] Yun-Jen Chiu, Van-Truong Pham, Thi-Thao Tran and Kuo-Kai Shyu,(2010) "Evaluation of active contour on medical inhomogeneous image segmentation", 3rd International Conference on Computer Science and Information Technology, DOI:10.1109/ICCSIT.2010.5564101,pp.311 – 314, 2010.

[5] Y.Ebrahimdoost, J.Dehmeshki, T.S.Ellis, M.Firoozbakht, A.Youannic and SD.Qanadli,(2010) "Medical Image Segmentation Using Active Contours and a Level Set Model: Application to Pulmonary Embolism (PE) Segmentation", IEEE Xplore, DOI: 10.1109/ICDS.2010.64, pp.269 – 273, 2010.

[6] Jieyu Cheng, Mingyue Ding and Xuming Zhang, (2011) "Geometric Active Contour Model and its Application to Carotid Plaque Detection", International Conference on Intelligent Computation and Bio-Medical Instrumentation, DOI:10.1109/ICBMI.2011.62, pp.101 – 104, 2011.

[7] Jinping Fan, Ruichun Wang, Shiguo Li and Chunxiao Zhang, (2012) "Automated cervical cell image segmentation using level set based active contour model", in 12th International Conference on Control Automation Robotics & Vision (ICARCV), DOI: 10.1109/ICARCV.2012.6485273, pp.877 – 882, 2012.

[8] Evgin Goceri, Mehmet Z. Unlu, Cuneyt Guzelis and Oguz Dicle, (2012) "An automatic level set based liver segmentation from MRI data sets", 3rd International Conference on Image Processing Theory, Tools and Applications (IPTA), DOI:10.1109/IPTA.2012.6469551, pp.192 – 197, 2012.

[9] Amit Satish Unde, V. A. Premprakash and Praveen Sankaran, (2012) "A novel edge detection approach on active contour for tumor segmentation", Students Conference on Engineering and Systems, DOI: 10.1109/SCES.2012.6199100, pp.1-6, 2012.

[10] Xie Li-Zhi, Tian Yun, Zhou Ming-Quan and Cao Rong-Fei,(2012) "Cerebrovascular segmentation based on region growing and level set algorithm", International Conference on Audio, Language and Image Processing, DOI: 10.1109/ICALIP.2012.6376733,pp.854 – 859, 2012.

[11] Hassan Mouhadjer, M. Mansour, M. Ouslim and B. Bouchiba, (2013) "Segmentation of human chromosome images using distance regularized level set evolution", 2nd International Conference on Advances in Biomedical Engineering, DOI:10.1109/ICABME.2013.6648 886, pp.215 – 218, 2013.

[12] Dinesh Reddy Byreddy and MV Raghunadh, (2014) "An application of geometric active contour in bio-medical engineering", International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA), DOI: 10.1109/CSCITA.2014.6839280, pp.322 – 326, 2014.

[13] Mengjuan Chen, Jianwei Li, Hanqing Zhao and Xiao Ma, (2014) "A local region-based level set algorithm for image segmentation", IEEE Workshop on Electronics, Computer and Applications, DOI: 10.1109/IWECA.2014.6845753, pp.844 – 847, 2014.

[14] Zhiwei Liu, Dongao Zhou, Qiang Lin and Jiayu Lin, (2015) "Active Contour Model Based on Local and Global Image Information", 7th International Conference on Intelligent Human-Machine Systems and Cybernetics, DOI: 10.1109/IHMSC.2015.63, pp.266 – 269, 2015.

[15] Snehal Kulkarni and Sanjay Shitole, (2016) "Region growing based novel approach for estimation of speckle in SAR images", IEEE Bombay Section Symposium (IBSS), DOI: 10.1109/IBSS.2016. 7940 200, pp.1-5, 2016.

[16] Harmeet Saini and Varsha Sahni, (2017) "Region Growing segmentation using de-noising algorithm for medical ultrasound images", 3rd International Conference on Computational Intelligence & Communication Technology (CICT), DOI: 10.1109/CIACT.2017 .7977 313, pp. 1 – 5, 2017.

[17] Soomro S, Munir A and Choi KN, "Hybrid two-stage active contour method with region and edge information for intensity inhomogeneous image segmentation", PLOS ONE 13(1): e0191827, [Online]. Available: https://doi.org/10.1371/ journal.pone. 019 1827.

[18] Shafiullah Soomro and Kwang Nam Choi, (2018) "Active Contours Based on An Anisotropic Diffusion", Digital Image Computing: Techniques and Applications (DICTA), DOI: 10.1109/DICTA.2018.8615767, pp. 1 – 6, 2018.

LOWER BOUND OF ZERO-VISIBILITY COPS AND ROBBER GAME ON THE CUBE GRID

Jiahui Wang and Farong Zhong

Department of Computer Science, Zhejiang Normal University, Jinhua, China.

ABSTRACT

Zero-visibility cops and robber game is a variant of the classical model. The robber is invisible in the zero-visibility game. In this paper, we study the zero-visibility cops and robber game on the cube grid. We first study a partition problem of the cube grid C_{n^3} *. Then we prove the lower bound on the zero-visibility cop number of the cube grid by using the results in the partition. We also show the lower bound is a quadratic polynomial about n.*

KEYWORDS

Cube Grid; Cops and robber; Partition; Zero-visibility Cop number

1. INTRODUCTION

Cops and robber game is an interesting pursuit-evasion game in graph theory, which was introduced by Nowakowski and Winkler [1] and Quilliot [2] in 1983. We consider the game of cops and robber on a graph G . There are two kinds of players: one robber and multiple cops. In each round, the cops move first and then the robber. The player can only move to the adjacent vertex or stay still. The robber is captured if a cop and the robber occupy the same vertex in a finite number of moves. Our goal is to determine the minimum number of cops to capture the robber.

Cops and robber game can be divided into three models based on visibility. In these three models, the robber knows the location of all cops at any time, but the cop's visibility of the robber is different. The first is that the cops know the perfect information of the robber, a lot of results can be found in [3,4,5,6]. The second is that the cops know the imperfect information of the robber. The visibility of the cops is limited to the distance between the cops and the robber within k range, so this model is called the k-visibility cops and robber game. In this case, related results can be found in [7,8].

Zero-visibility cops and robber game is the third model, in which the cops have no information about the location of the robber, the robber is invisible. The zero-visibility cop number of G, denoted $c_0(G)$, is the minimum number of cops needed to capture the invisible robber, and a strategy used $c_0(G)$ cops is called the optimal search strategy. It was first proposed by Tošić [9] in 1985. Zero-visibility model makes the problem more difficult, so there are not many results. The key to determine $c_0(G)$ is to find the lower bound. Yang [10] proposed a partition method, which can be applied to prove the lower bound of $c_0(G)$. this method only used for simple

David C. Wyld et al. (Eds): SIGI, CSTY, AI - 2020
pp. 21-29, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.102003

cartesian products in [10]. Cube grid is a three-dimensional graph, so we use this method on the cube grid. Given an integer k, and find a partition of the cube grid such that the size of the boundary of the smaller subset in the partition is at most k while the size of the subset is as large as possible [10]. This method can successfully help us determine the lower bound of $c_0(G_{n^3})$,

The rest of this paper is as follows. In section 2, we give the related work. In section 3, we give some definitions. In section 4, we study a partition problem of the cube grid. In section 5, we show the lower bound of $c_0(C_{n^3})$ is a quadratic polynomial about n, and the last section is a summary.

2. RELATED WORK

Tošić [9] studied the zero-visibility cop numbers of paths, cycles, complete graphs, and complete bipartite graphs. Tang [11] proposed a quadratic time algorithm to determine the zero-visibility cop number of a tree, then a linear algorithm of a tree was given by Dereniowskl [12]. In the above articles, only simple graphs were studied, and there is no better method to determine the zero-visibility cop number until Dereniowskl [13] studied the relationship between the monotonic zero-visibility cop number and the path width of a graph. Some scholars have calculated the path width of different graphs in [14], also the path width of the cube grid can be found in [15]. But how to determine the lower bound of $c_0(G)$ is also difficult. Yang [10] proposed to use the partition method to prove the lower bound of the cop number in a simple Cartesian product graph, and new results of this method to determine the zero-visibility cop number can be found in [16].

3. PRELIMINARIES

Consider the zero-visibility cops and robber game on a graph $G(V,E)$ with vertex set V and edge set E, which can be seen as a node-search problem. First, make all vertices contaminated, then the robber can only be on contaminated vertices. Thus, the search is successful if the givencop number and search strategy can make all vertices clean.

Let G and G' be two graphs, the cartesian product of G and G', denoted by $G \square G'$, is the graph in which $V(G \square G') = V(G) \times V(G')$, and two vertices (u,v) and (u',v') are adjacent in $G \square G'$ if and only if $u = u'$, v and v' are adjacent in G'; or $v = v'$, u and u' are adjacent in G. Let P_n be a path with n vertices. The square grid is the cartesian product of P_n and P_n, denoted by $C_{n^2}(n \geq 2)$. The cube grid is the cartesian product of C_{n^2} and P_n, denoted by $C_{n^3}(n \geq 2)$, which has n^3 vertices. we define the direction in the cubic grid, let $d : \{x, y, z\}$, see in Figure 1. In each direction, there are n copies of C_{n^2} and n^2 copies of P_n. C_{3^2} and C_{3^3} shown in Figure 1.

Let (V_1, V_2) be a partition of $G(V,E)$ if $V_1 \cup V_2 = V$ and $V_1 \cap V_2 = \phi$. Two vertices v_i and v_j are adjacent, denoted by $v_i \sim v_j$. Let V_B be the boundary of V_1 if $V_B \subseteq V_1$, $V_B = \{v_i \mid v_i \in V_1, \exists v_j \in V_2 \; and \; v_i \sim v_j\}$, $|V_B|$ to be largest. The vertex in V_1 not in V_B, denoted by

V_C, we have $V_C = V_1 - V_B$. In the zero-visibility cops and robber searching model, the contaminated and cleared vertices always changed after each round. Let $V_B^i (i \geq 1)$ be the boundary after the cops' move in round i, V_1^i, V_2^i be the set of the cleared and contaminated vertices after the robber' move in round i. (V_1^i, V_2^i) is a partition of G after the robber' move in round i.

Each direction contains all vertices of the cube grid, so we take the z direction as an example. Let $C_{n^2}^l (1 \leq l \leq n)$ be the l-th C_{n^2} in the z direction, $v_{l,t} (1 \leq t \leq n^2)$ be the t-th vertex in the $C_{n^2}^l$ (see in figure 1). Let r_l, p_l, q_l be the number of vertices in V_1, V_C, V_B of $C_{n^2}^l (1 \leq l \leq n)$ respectively, we have $r_l = p_l + q_l$.

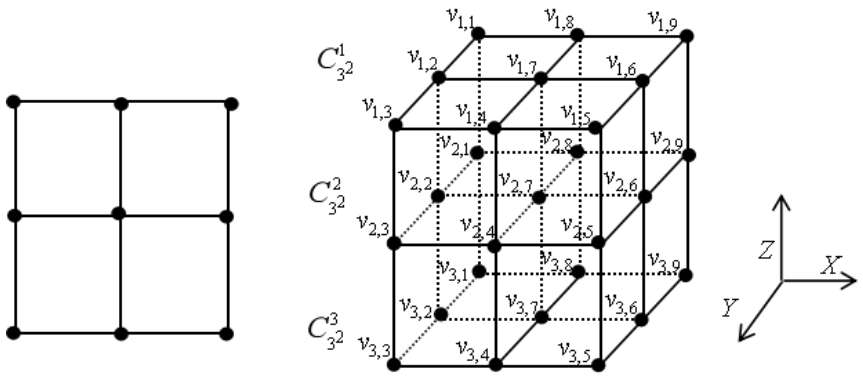

Figure 1. Square grid C_{3^2}, Cube grid C_{3^3}.

4. PARTITION OF THE CUBE GRID

Lemma 1. Let (V_1, V_2) be a partition of C_{n^2} for $n \geq 2$. If $|V_1| = \frac{m(m+1)}{2}$ and $1 \leq m \leq n$, then $|V_B| \geq m$.

Proof. The claim is true when $m = 1$. For the sake of contradiction, suppose that $|V_B| < m$ when $2 \leq m \leq n$. If there exist m rows or m columns that contain both vertices in V_1 and V_2, then $|V_B| \geq m$, which is a contradiction. Hence, there are at least $n-m+1$ rows and $n-m+1$ columns that contain vertices only in V_1 or V_2. Here we have two cases.

Case1: There exist $n-m+1$ rows and $n-m+1$ columns that contain vertices only in V_2. In this case, it can be seen that there exist a row and a column contain vertices only in V_2 of C_{m^2}. From theorem 1 in Yang [15], we have $|V_1| \leq \frac{m(m-1)}{2} < \frac{m(m+1)}{2}$.

Case2: There exist $n-m+1$ rows and $n-m+1$ columns that contain vertices only in V_1. Note that $|V_1|_{\min} = n^2 - |V_2|_{\max}$, the maximum of $|V_2|$ is equal to the maximum of $|V_C|$ in case 1,so consider case 1, we will show that when $|V_C|_{\max} = w$, $|V_B| = m-1$. Assume that $|V_B| = k < m-1$, $|V_C|_{\max} = w$. If there are $m-1$ rows or $m-1$ columns contain both vertices in V_1 and in V_2, then $|V_B| \geq m-1$, which is a contradiction. Thus, there are at least $n-m+2$ rows and $n-m+2$ columns that contain vertices only in V_2. Let v_i be a vertex in V_2, satisfied $\{v_i \mid v_i \sim v_j, v_j \in V_B\}$, then let $v_i \in V_1$, we can obtain a new partition of C_{n^2}. In a new partition, we have $|V_B| = k$, $|V_C| = w+1$; or $|V_B| = k+1, |V_C| = w$, which is a contradiction. Thus, when $|V_C|_{\max} = w$, we obtain $|V_B| = m-1$. Since $w = |V_1| - (m-1)$, $|V_1| \leq \frac{m(m-1)}{2}$, we have $w \leq \frac{m(m-1)}{2} - (m-1)$. Hence, in case 2, $|V_1|_{\min} = n^2 - [\frac{m(m-1)}{2} - (m-1)]$. Therefore, $|V_1| \geq \frac{m(m+3)}{2} - 1 > \frac{m(m+1)}{2} (m \geq 2)$.

See in figure 2.

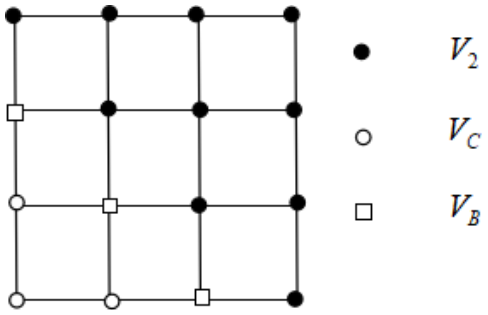

Figure 2. C_{4^2} with $m=3, |V_1|=6, |V_B|=3$.

Lemma 2. Let (V_1, V_2) be a partition of C_{n^3} for $n \geq 2$. If $\max\{r_l \mid 1 \leq l \leq n\} = m (m \leq \frac{n(n+1)}{2})$, then $|V_B| \geq m$.

Proof. We know that on a $C_{n^3} (n \geq 2)$, if $r_{l_1} \leq r_{l_2} \leq \frac{n(n+1)}{2}, (1 \leq l_1, l_2 \leq n)$, then $q_{l_1} \leq q_{l_2}$. So $|V_B| = \sum_{l=1}^{n} q_l$ is the minimum when $r_1 = m$ or $r_n = m$, this two cases are similar. Thus, take $r_n = m$ as an example, there will be $r_{l-1} \leq r_l$, $q_{l-1} \leq q_l (2 \leq l \leq n)$. Here are two cases.

Case1: $\min\{r_l \mid 1 \le l < n\} = 0$ .Suppose that $r_i = 0$, then let $v_{m,t}$ be a vertex in $C^m_{n^2}$ such that $v_{m,t} \in V_1$, there exist a path $v_{i,t}...v_{m,t}$ $(v_{i,t} \in C^i_{n^2}, v_{i,t} \in V_2)$ in z direction, and this path must contain a vertex in V_B. Since $r_i = m$, we have $|V_B| \ge m$. (see in figure 3)

Case2: $\min\{r_l \mid 1 \le l < n\} > 0$ .Suppose that $r_n = \frac{n(n+1)}{2}$, then $q_n \ge n$ from lemma 1. And let $q_n = n$ for minimum, we obtain $1 \le q_l \le n$, and we have $q_{l-1} \le q_l$ when $|V_B|$ is the minimum, so there exists one copy of C_{n^2} that contain one vertex in V_B. Since $r_n = m < \frac{n(n+1)}{2}$, there are at least two copies of C_{n^2} in which the number of boundaries equals 1 when $|V_B|$ is the minimum. Let $C^j_{n^2}$ be a copy of C_{n^2} in which $r_j = q_j = 1$. Suppose that there is no copy of C_{n^2} between $C^j_{n^2}$ and $C^n_{n^2}$ in which the boundary number equals 1, similar to (1), we have $|V_B| \ge m$. And there exist another copy of C_{n^2} whose boundary number equals 1, so we have $|V_B| \ge m+1$.(see in figure 4)

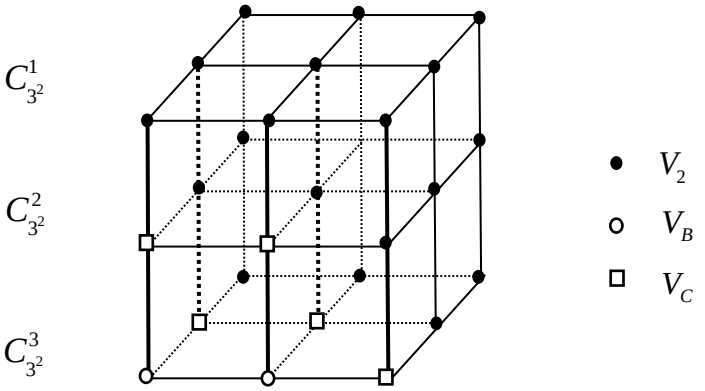

Figure 3. C_{3^3} with $i = 1, r_i = 0, m = 5$, and there exist 5 copies of P_3 (bold lines)which contain a vertex in V_1 and a vertex in V_2, so $|V_B| = 5$.

Theorem 1. Let (V_1, V_2) be a partition of C_{n^3} for $n \ge 2$. If $|V_B| < \frac{n(n+1)}{2} + 1$, and there exist $\frac{n(n-1)}{2}$ copies of P_n in which all vertices belong to V_2 in one direction, then $|V_1| \le \sum_{l=1}^{n} \frac{l(l+1)}{2}$.

Proof. Take z direction as an example. Since there exist $\frac{n(n-1)}{2}$ copies of P_n whose vertices all belong to V_2, then $r_l \le \frac{n(n+1)}{2} (1 \le l \le n)$. We need to find the maximum of $|V_1|$. Suppose that $\max\{r_l (1 \le l \le n)\} = \frac{n(n+1)}{2}$, from lemma2, we have $|V_B| \ge \frac{n(n+1)}{2}$ .And since

$|V_B|<\frac{n(n+1)}{2}+1$, we have $|V_B|=\frac{n(n+1)}{2}$. If $C_{n^2}^{i}, C_{n^2}^{i+1}.....C_{n^2}^{j}(1\le i\le j\le n)$ contain vertices in V_1, then $r_{k-1}=p_{k-1}+q_{k-1}\ge p_k(i+1\le k\le j)$. Hence, when $|V_B|=\frac{n(n+1)}{2}$, there will be $r_{k-1}=p_k(i+1\le k\le n)$, $p_i=0$, $r_i=q_i$, $r_n=\frac{n(n+1)}{2}$. And from lemma1, $q_n\ge n$. Since $r_{k-1}=p_k$, we have $r_{n-1}=p_n\le\frac{n(n+1)}{2}-n=\frac{n(n-1)}{2}$.Let $r_{n-1}=\frac{n(n-1)}{2}$ for maximum, also from lemma 1, there will be $q_{n-1}\ge n-1$, and so on. Thus, when $|V_1|$ is the maximum, we have $r_l=\frac{l(l+1)}{2}$, $q_l=l(1\le l\le n)$, satisfied $|V_B|=\sum_{l=1}^{n}q_l=\frac{n(n+1)}{2}<\frac{n(n+1)}{2}+1$. Therefore,

$|V_1|\le\sum_{l=1}^{n}\frac{l(l+1)}{2}$.(see in figure 4)

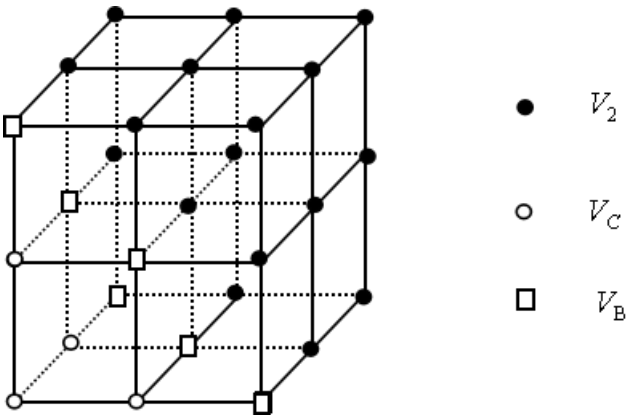

Figure 4. C_{3^3} with $|V_B|=6, |V_1|=10$.

5. LOWER BOUND OF ZERO-VISIBILITY COP NUMBER

Lemma 3([15]). Let G be a graph with $c_0(G)=k$. For any round $i(i\ge 1)$ of an optimal search strategy of G, if $|V_B^i|\ge 2k$, then $|V_1^i|\le|V_1^{i-1}|$.

Lemma4. For an optimal search strategy of C_{n^3}, if $|V_B^i|<\frac{n(n+1)}{2}+1$, then there will be $|V_1^i|\le\sum_{l=1}^{n}\frac{l(l+1)}{2}$ or $|V_1^i|\ge n^3-\sum_{l=1}^{n}\frac{l(l+1)}{2}+\frac{n(n+1)}{2}$.

Proof. After cops' move in round i, take z direction as an example, there are n^2 copies of P_n in z direction. If there are $\frac{n(n+1)}{2}+1$ copies of P_n that contain both cleared vertices and contaminated vertices, then $|V_B^i| \geq \frac{n(n+1)}{2}+1$, which is a contradiction. Hence, there are at least $\frac{n(n-1)}{2}$ copies of P_n in which all vertices are cleared or contaminated. Here we have two cases.

Case1: There exist $\frac{n(n-1)}{2}$ copies of P_n that contain only contaminated vertices. Let V_1, V_2 be the set of cleared vertices and contaminated vertices respectively after cops' move in round i. From theorem1, we have $|V_1| \leq \sum_{l=1}^{n} \frac{l(l+1)}{2}$. Therefore, $|V_1^i| \leq |V_1| \leq \sum_{l=1}^{n} \frac{l(l+1)}{2}$.

Case2: There exist $\frac{n(n-1)}{2}$ copies of P_n that contain only cleared vertices. The proof is similar to lemma1, consider case 1 again, we will show that when $|V_B^i| = \frac{n(n+1)}{2}$, $|V_1| - |V_B^i|$ is the maximum. Thus, in case2, the maximum number of contaminated vertices is $\sum_{l=1}^{n} \frac{l(l+1)}{2} - \frac{n(n+1)}{2}$. Therefore, we have $|V_1^i| \geq n^3 - \sum_{l=1}^{n} \frac{l(l+1)}{2} + \frac{n(n+1)}{2}$.

Theorem 2. For $n \geq 2$, $c_0(C_{n^3}) \geq \left\lfloor \frac{n(n+1)}{4} + \frac{3}{2} \right\rfloor$.

Proof. For the sake of contradiction, suppose that $c_0(C_{n^3}) \leq \left\lfloor \frac{n(n+1)}{4} + \frac{1}{2} \right\rfloor$, and consider a strategy that uses at most $\left\lfloor \frac{n(n+1)}{4} + \frac{1}{2} \right\rfloor$ cops, we will show that $|V_1^i| \leq \sum_{i=1}^{n} \frac{l(l+1)}{2} < n^3, (i \geq 0)$.

That means $\left\lfloor \frac{n(n+1)}{4} + \frac{1}{2} \right\rfloor$ cops can't make all vertices clean, which is a contradiction. When $i = 0$, we have $|V_1^0| \leq \left\lfloor \frac{n(n+1)}{4} + \frac{1}{2} \right\rfloor \leq \sum_{l=1}^{n} \frac{l(l+1)}{2}$; then assume that $|V_1^{i-1}| \leq \sum_{l=1}^{n} \frac{l(l+1)}{2}$, here we have two cases to prove $|V_1^i| \leq \sum_{l=1}^{n} \frac{l(l+1)}{2}$.

Case1: $|V_B^i| \geq \frac{n(n+1)}{2}+1$. From lemma 3, $|V_1^i| \leq |V_1^{i-1}|$. Therefore, $|V_1^i| \leq \sum_{l=1}^{n} \frac{l(l+1)}{2}$.

Case2: $|V_B^i| < \frac{n(n+1)}{2}+1$. Since $|V_1^{i-1}| \le \sum_{l=1}^{n}\frac{l(l+1)}{2}$, there are at most $\left\lfloor \frac{n(n+1)}{4}+\frac{1}{2}\right\rfloor + \sum_{l=1}^{n}\frac{l(l+1)}{2}$ cleared vertices after cops' move in round i. Also, since $n^3 - \sum_{l=1}^{n}\frac{l(l+1)}{2} + \frac{n(n+1)}{2} > \left\lfloor \frac{n(n+1)}{4}+\frac{1}{2}\right\rfloor + \sum_{l=1}^{n}\frac{l(l+1)}{2}$, we have $|V_1^i| \le \sum_{l=1}^{n}\frac{l(l+1)}{2}$ from lemma 4.

6. CONCLUSION

In this paper, we successfully apply the partition method to study zero-visibility cops and robber game on the cube grid and show the lower bound of $c_0(C_{n^3})$. Perhaps the lower bound obtained by this method is not the best, But we narrow the scope of finding the final solution. In the future, we will use the result in this paper to find the exact value of $c_0(C_{n^3})$ and also use this method on other three-dimensional graph.

REFERENCES

[1] R. Nowakowski and P. Winkler(1983). Vertex to vertex pursuit in a graph. Discrete Mathematics 43:235–239.

[2] Quilliot, A.(1983): Problèmes de jeux, de point fifixe, de connectivité et de représentation sur des graphes, des ensembles ordonnés et des hypergraphes. PhD thesis, Université de Paris VI

[3] Anthony Bonato. (2011). The game of cops and robbers on graphs. Student Mathematical Library, 276.

[4] Anthony Bonato, & Boting Yang. (2013). Graph searching and related problems. *Handbook of Combinatorial Optimization*, 1511-1558.

[5] Bonato, A. , & Macgillivray, G. . (2017). Characterizations and algorithms for generalized cops and robbers games.

[6] Brandt, S. , Emek, Y. , Uitto, J. , & Wattenhofer, R. . (2020). A tight lower bound for the capture time of the cops and robbers game. Theoretical Computer ence.

[7] Isler, V., Karnad, N.(2008): The role of information in the cop-robber game. Theoretical Computer Science 399, 179–190

[8] Clarke, N. E. , Cox, D. , Duffy, C. , Dyer, D. , & Messinger, M. E. . (2019). Limited visibility cops and robber. Discrete Applied Mathematics, 282.

[9] Tošić R.(1985), Vertex-to-Vertex Search in a Graph, Proceedings of the Sixth Yugoslav Seminar on Graph Theory , 233-237.

[10] Xue, Y. , Yang, B. , Zhong, F. , & Zilles, S. . (2019). A Partition Approach to Lower Bounds for Zero-Visibility Cops and Robber. Combinatorial Algorithms.

[11] Tang, A.(2004): Cops and robber with bounded visibility[D]. Master's thesis, Dalhousie University.

[12] Dereniowski, D. , Dyer, D. , Tifenbach, R. M. , & Yang, B. . (2014). The complexity of zero-visibility cops and robber.

[13] Dariusz Dereniowski, Danny Dyer, Ryan M. Tifenbach, & Boting Yang. (2013). Zero-visibility cops and robber game on a graph. *Lecture Notes in Computer Science, 7924*, 175-186.

[14] Ellis, J. , & Warren, R. . (2008). Lower bounds on the pathwidth of some grid-like graphs. Discrete Applied Mathematics.

[15] Otachi, Y. , & Suda, R. . (2011). Bandwidth and pathwidth of three-dimensional grids. *Discrete Mathematics, 311*(10), 881-887.

[16] Xue, Y. , Yang, B. , & Zilles, S. . (2019). New Results on the Zero-Visibility Cops and Robber Game. International Conference on Algorithmic Applications in Management. Springer, Cham.

AUTHORS

Jiahui Wang: an M.S. candidate in Computer Science at Zhejiang Normal University. Her research interests focuses on Graph searching.

Farong Zhong: a professor in Computer Science at Zhejiang Normal University. He received his B.S.degree in Computer Science from Shandong University in 1986, M.S. and Ph.D. degree in Computer Science from Shanghai Jiaotong University in 1994 and 2005 respectively. His research interests focuses on process calculus, Web services, and Graph searching.

PERFORMANCE ANALYSIS OF MACHINE LEARNING CLASSIFIERS FOR INTRUSION DETECTION USING UNSW-NB15 DATASET

Geeta Kocher[1] and Gulshan Kumar[2]

[1]Research Scholar, Department of Computational Sciences,
MRSPTU, Bathinda, Punjab, India
[2]Assosiate Professor, Department of Computer Applications, SBSSTC,
Ferozpur, Punjab, India

ABSTRACT

With the advancement of internet technology, the numbers of threats are also rising exponentially. To reduce the impact of these threats, researchers have proposed many solutions for intrusion detection. In the literature, various machine learning classifiers are trained on older datasets for intrusion detection which limits their detection accuracy. So, there is a need to train the machine learning classifiers on latest dataset. In this paper, UNSW-NB15, the latest dataset is used to train machine learning classifiers. On the basis of theoretical analysis, taxonomy is proposed in terms of lazy and eager learners. From this proposed taxonomy, K-Nearest Neighbors (KNN), Stochastic Gradient Descent (SGD), Decision Tree (DT), Random Forest (RF), Logistic Regression (LR) and Naïve Bayes (NB) classifiers are selected for training. The performance of these classifiers is tested in terms of Accuracy, Mean Squared Error (MSE), Precision, Recall, F1-Score, True Positive Rate (TPR) and False Positive Rate (FPR) on UNSW-NB15 dataset and comparative analysis of these machine learning classifiers is carried out. The experimental results show that RF classifier outperforms other classifiers.

KEYWORDS

Intrusion Detection System, Random Forest, KNN, UNSW-NB15, Machine Learning Classifiers

1. INTRODUCTION

Nowadays, to secure the confidential data from the eye of attackers is becoming a crucial and difficult task. The traditional methods like firewall and antivirus are not sufficient to tackle all types of attacks. So, there is a need for additional security along with traditional methods. Intrusion Detection System (IDS) play a significant role in this regard. It carefully keeps a track on the network traffic data and distinguishes whether the data is normal or attack.

An IDS is used to monitor the network traffic for detecting malicious activity. It can easily detect the attacks that are bypassed by the firewall. It continuously monitors the network, finds vulnerable parts of the network and communicates the administrator about intrusions [1]. It can be separated into two classes: anomaly detection and misuse detection. Misuse detection operates with prior prepared patterns of known attacks also called signatures [2]. It has high accuracy and low false alarm rates (FAR) but unable to detect novel attacks [3]. One of the solutions to address this problem is to regularly update the database which is not feasible and a costly process. So, anomaly detection techniques came into existence. Anomaly Detection deals with profiling user

David C. Wyld et al. (Eds): SIGI, CSTY, AI - 2020
pp. 31-40, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.102004

behaviour [4]. In this approach, a certain model of user normal activity is defined, and any deviation from this model is known as anomalous. Fig. 1 shows the diagram of IDS.

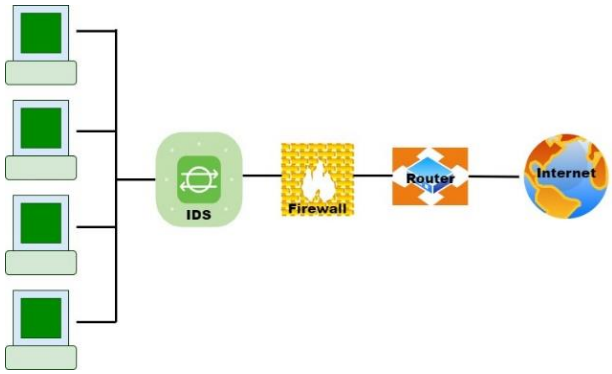

Fig 1: Intrusion Detection System

In literature, different types of machine learning (ML) classifiers are used for intrusion detection. From the literature, it is found that there is limited work done on comparative analysis of ML classifiers. Hence, the motive of this paper is to find performance comparison of several ML classifiers using recent dataset for intrusion detection. The structure of the paper is divided into seven sections. Section 2 gives a brief literature survey related to this research work. In section 3, taxonomy of the classifiers is discussed. A brief introduction about the dataset used for experimental work is described in Section 4. In Section 5, a methodology to pre-process the dataset is presented. Experimental work is shown in Section 6 and Section 7 gives Conclusion and Future scope.

2. RELATED WORK

This section provides the literature survey on the ML classifiers. The main motive of this section is to give an overview of the research work done in the field of intrusion detection. It is found from the literature that researchers have put a lot of efforts on ML classifiers and some of their contributions are described below:

Narudin et. al. (2014) [5] described an evaluation using ML classifiers namely RF, J-48, Multilayer Perceptron (MLP), NB and KNN to detect mobile malware using two datasets namely MalGenome and Private. Weka Tool was used for the evaluation. The performance metrics namely TPR, FPR, precision, recall and f-measure were used to validate the performance of ML classifiers. The accuracy obtained by RF Classifier is 99.99% during experimental work on MalGenome dataset. The author has suggested to improve the results by using selected features for future work.

Belavagi & Muniyal, (2016) [6] designed a Network Intrusion Detection System (NIDS) with the various supervised machine learning classifiers. NSL-KDD dataset was used to check the performance of various classifiers. The result shows that RF classifier outperforms other classifiers. It results in lowest FPR and a highest TPR and accuracy obtained is 99%. But still, there is a need for classifiers that can be used for the multiclass classification.

Ashfaq et al, (2017) [7] described a semi-supervised learning (SSL) approach based on novel fuzziness. In order to improve the classifier performance, it utilizes unlabelled samples along with a supervised learning algorithm. NSL-KDD dataset was used for evaluation of this model.

The limitation of this model was that its performance was studied only for the Binary classification task.

Yaseen et al, (2017) [8] described a multilevel hybrid intrusion detection model using Support Vector Machine (SVM) and Extreme Learning Machine (EVM). The evaluation was done on KDD 99 dataset. The accuracy obtained was 95.75% and shorter training time in this proposed model. This technique is better only for known attacks and for novel attacks, efficient classifiers are required.

Aljumah, (2017) [9] described a trained algorithm to detect DDoS attacks which was based on Artificial Neural Network (ANN). ANN shows 92% accuracy when it was trained with older data sets and when the system is trained with updated datasets, the accuracy obtained was 98%. The accuracy of the ANN model depends upon the dataset. So, there is a need for the up-to-date and balanced dataset.

Roshan et al., (2018) [10] discussed an adaptive design of IDS based on Extreme Learning Machines (ELM). The NSL-KDD dataset was applied for the evaluation. It was found that it can detect novel attacks along with known attacks with an acceptable rate of detection and false positives.

Ali et al., (2018) [11] proposed a PSO-FLN classifier for intrusion detection. The benchmark dataset KDD99 was used to validate the results. PSO-FLN has outperformed ELM and FLN classifiers in terms of accuracy. But for some classes like R2L, it does not show accurate results.
From the literature survey, it is concluded that most of the researches have been validated using older datasets. These datasets lack novel attacks and contains imbalanced network audit data. Non-uniform distribution of data may lead to biased training of ML classifiers and this problem needs to be resolved. The new dataset can be used to detect novel attacks. The RF classifier shows better results as compared to other classifiers. A lot of work is done on binary classification and still there is a need to work more on multi-classification.

3. TAXONOMY OF CLASSIFIERS

The classifiers are divided into two learning methods i.e. lazy learners and eager learners [12-15]. The taxonomy of classifiers is proposed based on theoretical analysis and is shown in Fig. 2.

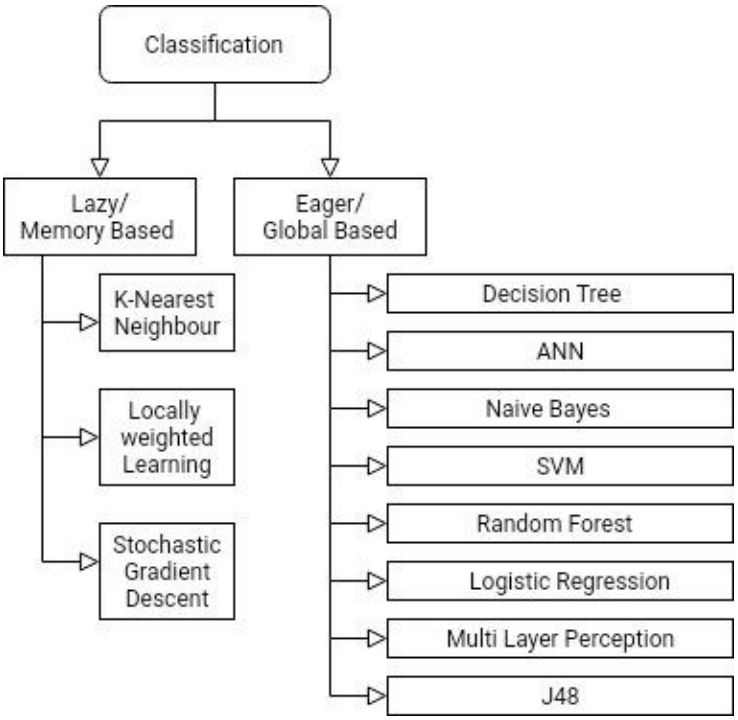

Fig 2: Taxonomy of classifier

The lazy learners can store examples and solve multiple problems with these examples. These learners adapt automatically to changes in the problem domain and easy to maintain. But the limitation of these learners is that they stored the same kind of examples many times and due to this require high memory and time-consuming learners. Eager learners firstly build a classification model on given training data and then perform a classification. These learners take more time for learning and less time for classifying the data.

From the above taxonomy of classifiers, KNN, SGD, DT, NB, RF and LR are used for experimental work in this paper. The description of these classifiers which are explored for experimental work is given below:

3.1. Lazy Learners

These learners use the training data for storage and wait for testing data to appear. KNN, Locally weighted learning (LWL) and SGD are examples of Lazy learners.

3.1.1. K- Nearest Neighbour

It is a lazy learning algorithm that firstly stores all the training data. At the time of classification, it uses this data and tries to find the similarities between the new data and the available data. It places the new data in the category that is most similar to the available data. It is based on the Euclidean distance [16]. The test data is allotted to the class of its K nearest neighbours. As you increase the value of K, accuracy might increase. It can be used for both regression and classification but is often used for classification problems.

3.2. Eager Learners

Eager learners take a long time for training and less time for predicting. DT, NB, LR, SVM, RF, MLP and J48 are examples of eager learners.

3.2.1. Decision Tree

It is a popular and powerful tool for prediction and classification. The structure of DT is like a tree structure in which each internal node represents a test on an attribute, each branch interprets a result of the test, and each leaf node shows a class label. DT perform classification without requiring much computation and able to handle both categorical and continuous features. This tree structure is computationally expensive to train and shows errors in multi-classification problems [17].

3.2.2. Logistic Regression

It is applied to solve both binary class and multiclass classification problems. The probability of occurrence of an event is predicted by giving fitting data to Logistic function. The output of this function lies between 0 and 1. The middle value i.e. 0.5 is considered as the threshold between class 1 and class 0. The output greater than 0.5 is considered as class 1 and if output is below 0.5, then it is considered as class 0[6].

3.2.3. Random Forest

It is proposed by Breiman in 2001. This method is based on the proximity search and can be used both for regression and classification. It is a decision tree-based classifier. In this technique, random samples are used to create decision trees, and then prediction is done from each tree and the best solution is found out by voting [16]. Random forest has many applications like image classification, feature selection and recommendation engines.

3.2.4. Naive Bayes

It is a classification algorithm used both for two class and multi-class classification problems. It assumes that probabilities of every feature belonging to each class are used for prediction [6]. It also assumes that the probability of every feature belonging to a given class value is independent from other features. For the known value of the feature, probability is known as conditional probabilities. Prediction can be attained by calculating the instance probabilities of each class and by selecting the class value of highest probability [12].

4. DATASET USED

The benchmark datasets used in the literature are older datasets and contains repeated records due to which ML classifiers give unfair results. So, selected ML classifiers are tested using UNSW-NB15 dataset which is novel dataset [18]. This dataset is composed of 49 attributes including a class label and contains 25,40,044 labelled instances, each being labelled either normal or attack. A detailed description about the features is given in Table 1.Table 2 gives the details of attacks.

Table 1: Description of the attributes of UNSW-NB15 dataset

S.No.	Type of attributes	Name of attributes	Sequence No.
1	Flow	Script, Sport, Dstip. Dsport, Proto	1-5
2	Basic	State, Dur, Sbytes, Dbytes, Sttl, Dttl, Sloss, Dloss, Service, Sload, Dload, Spkts, Dpkts	6-18
3	Content	Swin, Dwin, Stepb, Dtcpb, Smeansz, Dmeansz, trans_depth, res_bdy_len	19-26
4	Time	Sjit, Djit, Stime , Ltime, Sintpkt, Dintpkt, Tcprtt, Synack, Ackdat	27-35
5	General Purpose	is_sm_ips_ports, ct_state_ttl, ct_flw_http_mthd,is_ftp_login, ct_ftp_cmd	36-40
6	Connection	ct_srv_src , ct_srv_dst , ct_dst_ltm, ct_src_ ltm ,ct_src_dport_ltm,ct_dst_sport_ltm,ct_dst_src_ltm	41-47
7	Labelled	attack_cat, Label	48-49

When UNSW-NB15 dataset is used for evaluation, out of 49 attributes, we got 45 attributesonly. The four ID attributes are combined together to make a single attribute as ID from flow attribute category and two attributes of time category (Stime and Ltime) are combined together in one attribute known as Rate. The 42 attributes of UNSW-NB15 dataset are used to carry out experiment on ML classifiers. The dropout attributes are ID, Duration and attack_cat.

Table 2: Types of attacks in dataset

Type	**Whole**	**Training**
	No. of Records	**No. of Records**
Normal	2218761	56000
Fuzzers	24246	18184
Analysis	2677	2000
Backdoors	2329	1746
DOS	16353	12264
Exploits	44525	33393
Generic	215481	40000
Reconnaissance	13987	10491
ShellCode	1511	1133
Worms	174	130
		175341

5. METHODOLOGY

The pre-processing steps are shown in Fig.3 and Fig. 4 shows methodology used. In pre-processing, first of all the null values present in the dataset are handled. The categorical data is converted into numerical form with the help of label encoder. Then one hot encoder is used to break the relation between the values obtained through label encoder.

After this, the pre-processed data is separated as training and testing. The KNN, LR, NB, SGD, DT and RF classifiers are used to construct the models. Then the prediction of labels of test data is done using these models. A comparison is carried out between actual labels and predicted labels. The performance metrics used to evaluate the models are accuracy, precision, mean square error, recall, f1-score, TPR and FPR.

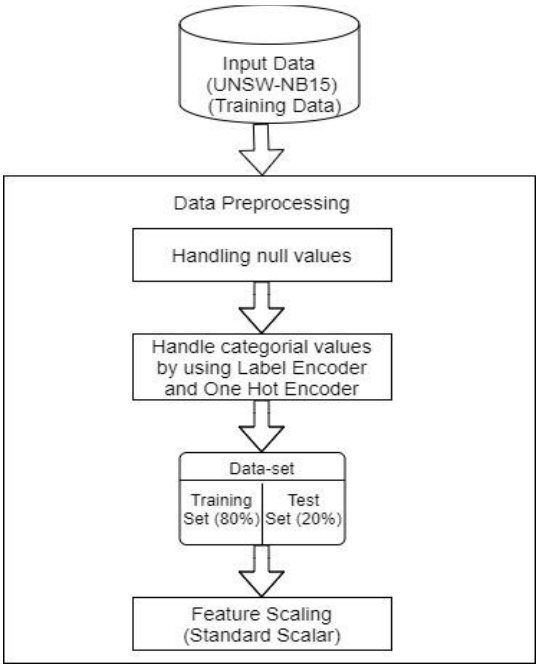

Fig. 3 Pre-processing steps on the UNSW-NB15 dataset

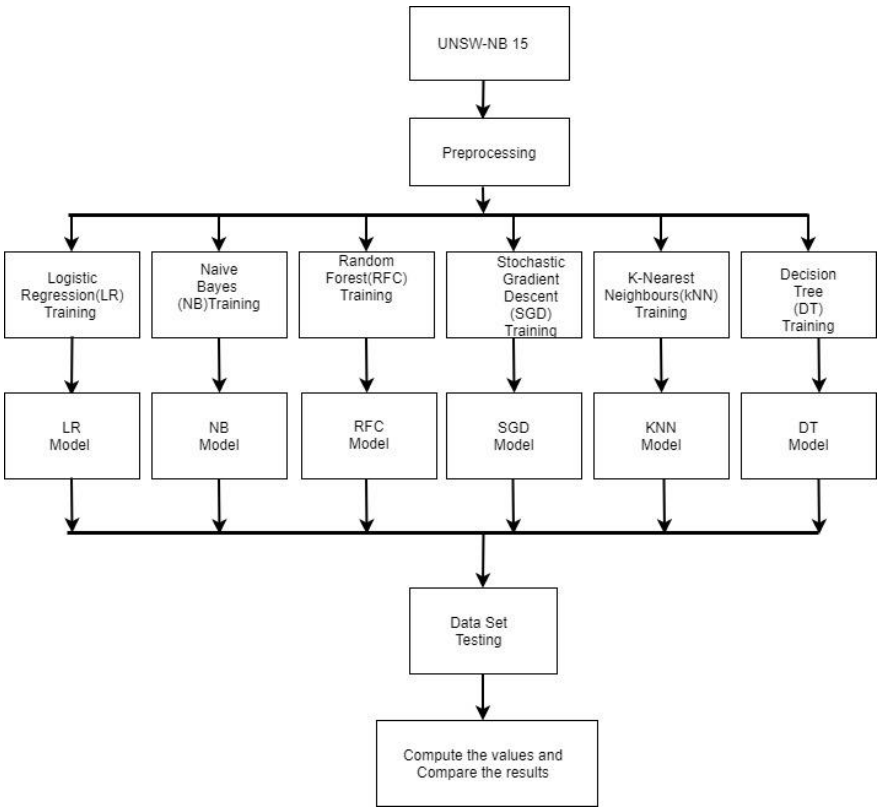

Fig. 4 Methodology

The procedural steps to construct the models are given below:

1. Start with pre-processing of dataset.
2. Divide dataset in to two parts i.e. training and testing.
3. Construct the classifier model using training data for KNN, LR, NB, RF, SGD and DT.
4. Take the test data
5. Testing of classifier models using training data
6. Calculate and compare Accuracy, Recall, Precision, F1-Score and Mean Squared Error for the selected models.

6. EXPERIMENTAL WORK

The selected ML classifiers namely LR, NB, RF, SGD, KNN and DT are tested on UNSW-NB15 dataset, the novel dataset for intrusion detection. The experimental work is done on Intel Core (TM) i3-1005G1 CPU @1.20 GHz, 4GB RAM using Python. After performing pre-processing steps, dataset is divided into two parts: training and testing. Then six classifiers are used for training as shown in Fig. 4 and performance is evaluated on the basis of several parameters as shown in Table 3. Fig. 5 shows the pictorial representation of accuracy of selected classifiers.

It can be observed from the results shown in Table 3 that the RF classifier out performs the other methods in terms of accuracy 95.43%, FPR 0.08 and mean squared error 0.046 whereas the NB shows the highest mean squared error 0.519 and lowest accuracy 48.03% in the selected group of classifiers.

Table 3: Performance comparison of selected classifiers using UNSW-NB15 dataset with train test split method

Classifier	Accuracy	Precision	Recall	F1-Score	Mean Squared Error	TPR	FPR
LR	93.23	0.92	0.99	0.95	0.068	0.99	0.19
NB	48.03	1.00	0.23	0.38	0.519	0.23	0
RF	95.43	0.96	0.97	0.97	0.046	0.97	0.08
SGD	93.29	0.91	1.00	0.95	0.067	0.99	0.21
KNN	93.71	0.94	0.96	0.95	0.063	0.96	0.12
DT	94.20	0.93	0.98	0.96	0.058	0.98	0.14

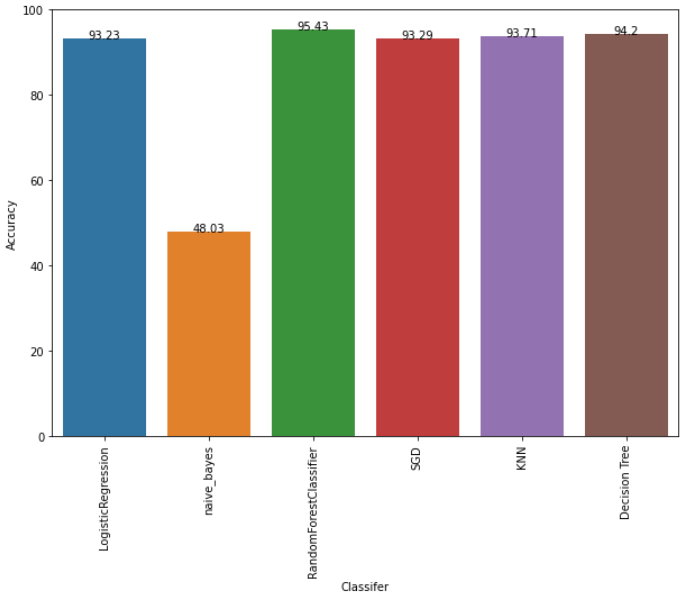

Fig. 5: Accuracy of selected classifiers using UNSW-NB15 dataset with train test Split method

7. CONCLUSION AND FUTURE SCOPE

The taxonomy of classifiers is proposed in terms of lazy and eager learners. The experimental work has been carried out to evaluate the performance of the selected ML classifiers based on proposed taxonomy namely KNN, LR, NB, DT, SGD and RF for detection of intrusion. These classifiers are tested on UNSW-NB15 data-set. The classifiers are compared on the basis of precision, MSE, recall, F1-Score, accuracy, TPR and FPR. The results show that RF classifier is better than other classifiers on UNSW-dataset using selected parameters. The accuracy of RF classifier comes out to be 95.43%. In future, this work can be extended for selective attributes and multiclass classification for detection of intrusion.

REFERENCES

1. Sarmah, A. (2001). Intrusion detection systems: Definition, need and challenges.
2. Omer, K. A. A., & Awn, F. A. (2015). Performance Evaluation of Intrusion Detection Systems using ANN. *Egyptian Computer Science Journal*, *39*(4).
3. Diro, A. A., &Chilamkurti, N. (2018). Distributed attack detection scheme using deep learning approach for Internet of Things. *Future Generation Computer Systems*, *82*, 761-768.
4. Khan, J. A., & Jain, N. (2016). A survey on intrusion detection systems and classification techniques. *Int. J. Sci. Res. Sci., Eng. Technol.*, *2*(5), 202-208.
5. Narudin, F. A., Feizollah, A., Anuar, N. B., &Gani, A. (2016). Evaluation of machine learning classifiers for mobile malware detection. *Soft Computing*, *20*(1), 343-357.
6. Belavagi, M. C., &Muniyal, B. (2016). Performance evaluation of supervised machine learning algorithms for intrusion detection. *Procedia Computer Science*, *89*, 117-123.
7. Ashfaq, R. A. R., Wang, X. Z., Huang, J. Z., Abbas, H., & He, Y. L. (2017). Fuzziness based semi-supervised learning approach for intrusion detection system. *Information Sciences*, *378*, 484-497.
8. Al-Yaseen, W. L., Othman, Z. A., &Nazri, M. Z. A. (2017). Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system. *Expert Systems with Applications*, *67*, 296-303.
9. Aljumah, A. (2017). Detection of Distributed Denial of Service Attacks Using Artificial Neural Networks. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *8*(8).

10. Roshan, S., Miche, Y., Akusok, A., & Lendasse, A. (2018). Adaptive and online network intrusion detection system using clustering and Extreme Learning Machines. *Journal of the Franklin Institute*, *355*(4), 1752-1779.
11. Ali, M. H., Al Mohammed, B. A. D., Ismail, A., &Zolkipli, M. F. (2018). A new intrusion detection system based on Fast Learning Network and Particle swarm optimization. *IEEE Access*, *6*, 20255-20261.
12. Bhavani, D.D., Vasavi, A. & Keshava P.T. (2016). Machine Learning: A Critical Review of Classification Techniques. IJARCCE, 5(3), 22-28.
13. Wei, C. C. (2015). Comparing lazy and eager learning models for water level forecasting in river-reservoir basins of inundation regions. Environmental Modelling & Software, 63, 137-155.
14. Rafatirad, S., & Heidari, M. (2018). An Exhaustive Analysis of Lazy vs. Eager Learning Methods for Real-Estate Property Investment.
15. https://towardsdatascience.com/machine-learning-classifiers-a5cc4e1b0623
16. Narudin, F. A., Feizollah, A., Anuar, N. B., &Gani, A. (2016). Evaluation of machine learning classifiers for mobile malware detection. Soft Computing, 20(1), 343-357.
17. https://www.geeksforgeeks.org/decision-tree/
18. Moustafa, N., & Slay, J. (2016). The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. Information Security Journal: A Global Perspective, 25(1-3), 18-31

AUTHORS

Ms. Geeta Kocher

She is MCA, M.Tech, M.Phil. She is pursuing Ph.D in the field of Artificial Intelligence-Deep learning. She has published more than 15 papers in various conferences and journals. She has more than 16 years experience in teaching.

Dr.Gulshan Kumar

Dr. Gulshan Kumar has received his MCA degree from Guru Nanak Dev University Amritsar (Punjab) India in 2001, and M.Tech. Degree in Computer Science & Engineering from JRN Rajasthan Vidyapeeth Deemed University, Udaipur (Rajasthan)-India, in 2009. He got his Ph.D. from Punjab Technical University, Jalandhar (Punjab)-India. He has 17 year of teaching experience. He has 56 international and national publications to his name. Currently, he is working as Associate Professor in Computer Applications department at Shaheed Bhagat Singh State Technical Campus, Ferozepur (Punjab)-India. He has supervised 06 M. Tech. students for their final thesis, students for projects MCA and supervising 02 PhD research scholar. His current research interests involve Artificial Intelligence, Network Security, Machine Learning and Databases.

MULTI IMAGE STEGANOGRAPHY USING DISTRIBUTED LSB ALGORITHM AND SECRET TEXT RECOVERY ON STEGO IMAGE CORRUPTION

Jagan Raj Jayapandiyan[1], C. Kavitha[2] and K. Sakthivel[3]

[1]Department of Computer Science, Periyar University,
Salem, Tamil Nadu, India
[2]Department of Computer Science, Thiruvalluvar Govt. Arts College,
Rasipuram, Tamil Nadu, India
[3]Department of Computer Science and Engineering, K. S. Rangasamy
College of Technology, Namakkal, Tamil Nadu, India

ABSTRACT

In this proposed research work, an attempt has been made to use multiple image files for steganography encoding along with the capability of secret text recovery in the event of any image corruption during the transit. This algorithm is effective on the security factor of secret image since the embedded checksum will validate for any unauthorized users or intruders attempt to corrupt the picture in any aspect. If any of the stego image underwent any steganalysis or MiM attack, then this proposed algorithm can effectively regenerate the content of one stego image using other intact stego images received in the receiving end.

KEYWORDS

Steganography, Multi-cover image, secret message recovery.

1. INTRODUCTION

Steganography is a science of concealing a file, document, image or video inside another file, message, image or video, which continues to be an extremely flexible and powerful way to disguise or cover information in plain sight. Using steganography, there are several ways to hide records. The most popular technique is to insert data into digital images. We all know that digital images mean that there are many megabytes of pixel data. It allows space in the digital file for someone to embed steganographic secret data. A good programmer can alter the Least Significant Bits (LSB) of any media file with the use of steganographic applications and embeds a malicious code in the digital picture.

2. STEGANOGRAPHY

The first recorded uses of steganography can be traced back to 440 BC in Greece, when Herodotus mentions two examples in his Histories.[1] Histiaeus sent a message to his vassal, Aristagoras, by shaving the head of his most trusted servant, "marking" the message onto his scalp, then sending him on his way once his hair had regrown, with the instruction, "When thou art come to Miletus, bid Aristagoras shave thy head, and look thereon." Additionally, Demaratus

David C. Wyld et al. (Eds): SIGI, CSTY, AI - 2020
pp. 41-50, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.102005

sent a warning about a forthcoming attack to Greece by writing it directly on the wooden backing of a wax tablet before applying its beeswax surface. Wax tablets were in common use then as reusable writing surfaces, sometimes used for shorthand.

The second story [2] also came from Herodotus, which claims that a soldier named Demeratus needed to send a message to Sparta that Xerxes intended to invade Greece. Back then, the writing medium was text written on wax-covered tablets. Demeratus removed the wax from the tablet, wrote the secret message on the underlying wood, recovered the tablet with wax to make it appear as a blank tablet and finally sent the document without being detected. Romans used invisible inks, which were based on natural substances such as fruit juices and milk. This was accomplished by heating the hidden text, thus revealing its contents.

3. TYPES OF STEGANOGRAPHY

Based on the type of cover file being used in the steganography technique, various types of steganography methods are as follows.

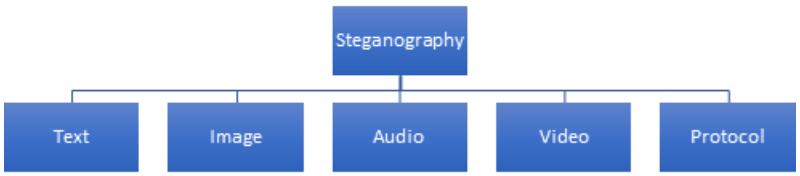

Figure 1. Type of Steganography

3.1. Text Steganography

The cover file used in the text steganography method would be in text format and the hidden message contained in the cover file would also be primarily text style. The embedding technique for text steganography is based on the number of characters, white spaces, capital letters, as used in the Morse language code in radio communication.

3.2. Image Steganography

Image steganography is a tool used to conceal a hidden message by taking the cover object as the image file. Graphic digital images are widely used as cover source in this steganography, and this cover file helps the user to embed a large volume of bits. The primary advantage of image steganography is that the attention of an intruder is not attracted by the cover image.

3.3. Audio Steganography

Audio steganography is a practice used to relay secret information in an imperceptible way by manipulating an audio signal. It is the science of hiding in a host message with any hidden text or audio content. The functions of the host message before steganography and the stego message after steganography are very identical. A more complicated method is embedding hidden messages in optical sound. Varieties of methods have been created for embedding information into digital audio.

3.4. Video Steganography

In multiple data hiding technology, video steganography is becoming a significant research field, which has become a promising technique. This is not only for the security necessity of secret message transmission becoming tighter, but video file also has enormous amount of data stream to leverage. Video steganography is broken down into three groups as per the embedded location of the hidden message: intra-embedding, pre-embedding and post-embedding [3]. Intra-embedding techniques are classified according to the phases of video encoding, such as intra-prediction, motion vectors, interpolation of pixels, coefficients of transformation. On raw footage, pre-embedding methods are manipulated, which can be divided into spatial domains and converted. Post-embedding strategies rely mostly on bitstreams, meaning that the operation of embedding and removing video steganography is all manipulated on the compressed bit stream.

3.5. Protocol Steganography

The Protocol Steganography is a modern solution for data hiding, which are popular in recent days. The network layer protocol of the TCP/IP (Transmission Control Protocol/Internet Protocol) suite is used for data hiding in this steganography and not limited exclusively to network protocols. For data hiding, covert channels are used in the network layer of the OSI architecture. Covert channels bypass the network system's security protocols. The intention is either used to steal information or use the network protocol to exchange hidden messages over a network.Example protocols used in the protocol steganography are TCP, IPv4 (Internet Protocol version 4), NFS (Network File Sharing), CIFS (Common Internet File System) etc.,

4. STEGANOGRAPHY PHASES

In order to complete the hidden message exchanging process from sender to receiver, every Steganography algorithm must come through various stages.

Sender: The prime objective of the sender is to embed the hidden message in the stego-medium and transmit it through the channel of communication.

Communication channel: A physical or wireless medium that holds an encoded cover picture across the network or some other distribution medium with a hidden message. The embedding strategy in the middle attacks should be sufficiently advanced to secure the hidden message for all potential intrusion.

Receiver: In this steganography process, it is the last stage where the cover medium is retrieved and extracted to see if the hidden text that was sent over the communication channel.

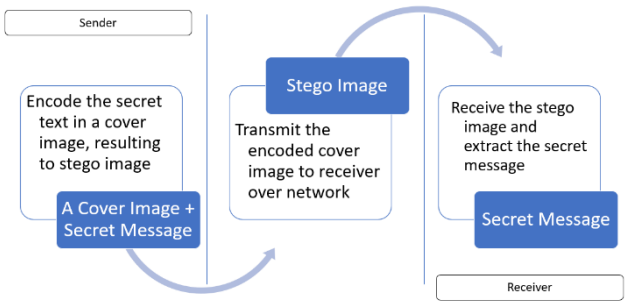

Figure 2. Phases of Steganography

5. RELATED WORKS

A succinct review based on the study of these papers related to our work is as follows. J. Homg et al. [4] and M. A. Hameed rt al [5] described several image steganography techniques in spatial domain. Along with existing techniques like LSB, layout management schemes and replacing only l's or only zero's, some more methods like replacing intermediate bit, raster scan principle, color-based data hiding and shape-based data hiding are also proposed. M. C. Kasapbasi et al. [6] and S. D. R. I. Moses [7] developed an improved method for image-based steganography using LSB technique. All these techniques are primarily focussed on the LSB steganography optimization and the steganography operation happens in one cover image. Though a high capacity focus is made on the research contributions in [4] to [7], these algorithms lack the ability to withstand the Steganalysis or man in the middle (MiM) attack on stego images during transmission with the intention of disrupting the transmission or acquiring the secret text that is in transit.

6. EXISTING STEGANOGRAPHY MODEL – LSB

A typical Steganography system consists of following elements.

- Cover Object (C)
- Secret Message (M)
- Stego Object (S)
-

6.1. Cover Object

The cover objects in Steganography are those in which we are hiding secret messages. The cover object can be any digital files such as photos, audio, writing, images. The cover object that is most used is an image file to hide information. Most of the times the cover image stays as single file in a steganography cycle.

6.2. Secret Message

The actual hidden message in Steganography process, which has to be hidden in the cover object. It is important that the hidden message does not cause any visible quality degradation to the cover object.

6.3. Stego Object

After hiding the secret in a cover object, now the object is called as stego object. then, the stego object is transferred over public post or transferred over an email to the receiving end to complete the cycle.

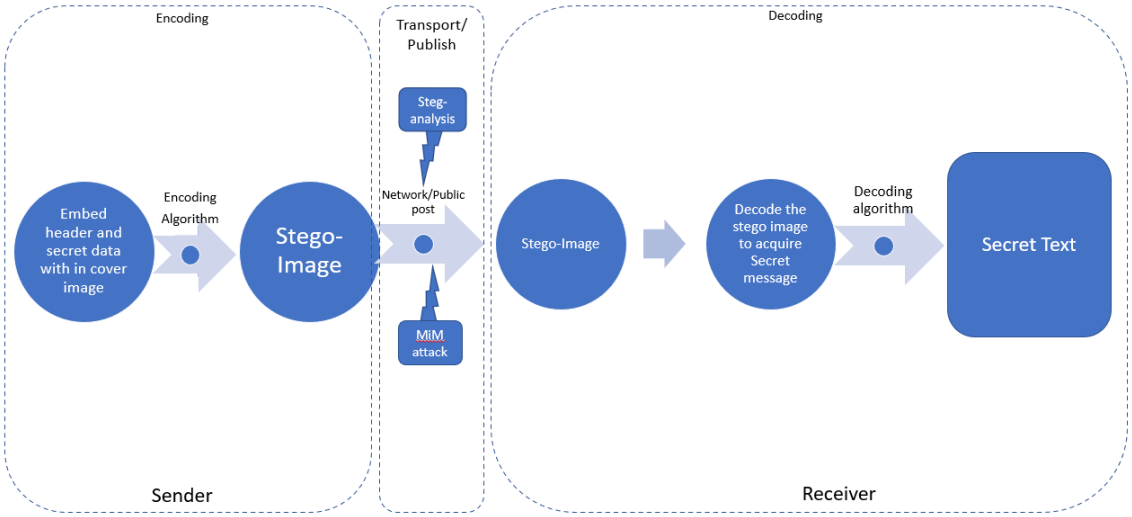

Figure 3. LSB Steganography model

7. PROPOSED MULTI-IMAGE COVER OBJECT MODEL

The proposed multi image cover model enables the receiver to send the secret text. This theoretical algorithm operates on the spatial domain of image steganography and places emphasis on maximizing the security of hidden message.

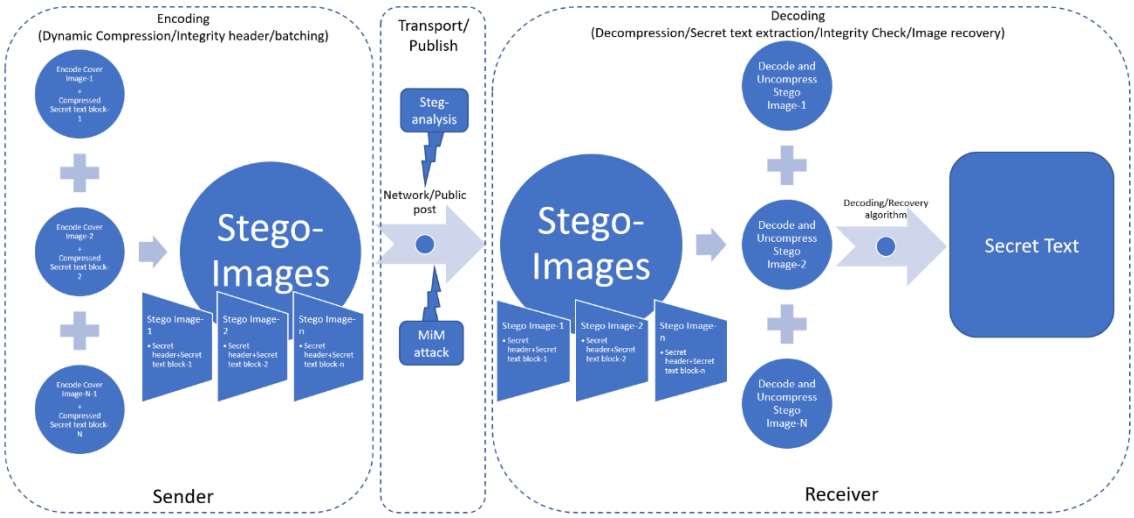

Figure 4. Proposed Steganography model

7.1. Embedding Algorithm

As part of secret message embedding in given cover image first the algorithm reads the height (X) and width (Y) of the cover image and collects meta data of a secret message like number of words characters. The primary inputs to the algorithm are cover image list (C), secret message (M) and number of least significant bit (k) that must be used during the embedding process.

Algorithm-1: Multi cover image embedding model
Input: Cover images, Message and Random seed
Output: Stego image

procedure eLSB_embed(C, M, k)
Read cover images, C
Read secret message, M
$X \leftarrow Height\ of\ the\ cover\ image, C$
$Y \leftarrow Width\ of\ the\ cover\ image, C$
$W \leftarrow Number\ of\ words\ in\ the\ message, M$
$L \leftarrow Number\ of\ characters\ in\ the\ message, M$
$E \leftarrow Equally\ divided\ secret\ message\ based\ on\ n(c)$
$T \leftarrow String\ vector\ of\ words\ from\ the\ secret\ message, M$
$S \leftarrow Secret\ message\ vector\ in\ binary\ form, T$
$Initialize, S \leftarrow [\,]$
$for\ w \leftarrow 1\ to\ W\ in\ steps\ of\ 1\ do$
$S_{[w]} \leftarrow binary((T_{[w]}))$
end for
$H \leftarrow binary\ (StegoHeader(S))$
$S \leftarrow H + S$
$for\ e \leftarrow 1\ to\ E\ in\ steps\ of\ 1\ do$
$for\ i \leftarrow 1\ to\ Y\ in\ steps\ of\ 1\ do$
$for\ j \leftarrow 1\ to\ X\ in\ steps\ of\ 1\ do$
$for\ x \leftarrow 1\ to\ 8\ in\ steps\ of\ 1\ do$
$rb = resetFromNthBit(k)_b$
$$C = \sum_{l=(8-k)}^{8} (c_{[i][j][x]} \& rb)$$
$$C = \sum_{l=(8-k)}^{8} (c_{[i][j][x]} \mid s_{[c++]})$$
end for
end for
end for
end for
return list C, the secret text embedded stego images
end procedure

7.2. Extraction Algorithm

Extraction algorithm is similar to compression procedure, but the steps are orderly reversed to obtain original secret message from stego-image.

Algorithm-2: Multi cover image extraction algorithm
Input: Stego image
Output: Secret message

procedure_eLSB_extract(C)
Read Stego images, C
$H \leftarrow extracted\ header\ from\ stego\ images, C$

$X \leftarrow Height\ of\ the\ cover\ image\ from\ header, H$
$Y \leftarrow Width\ of\ the\ cover\ image\ from\ header, H$
$k \leftarrow number\ of\ LSB\ used\ from\ header\ data, H$
$l \leftarrow Length\ of\ secret\ message\ in\ bytes\ from\ header\ data, H$
$E \leftarrow Equally\ divided\ secret\ message\ based\ on\ n(c), from\ header\ data, H$
$Initialize, h \leftarrow 64$
$for\ i \leftarrow 1\ to\ Y\ in\ steps\ of\ 1\ do$
$for\ e \leftarrow 1\ to\ E\ in\ steps\ of\ 1\ do$
$for\ j \leftarrow 1\ to\ X\ in\ steps\ of\ 1\ do$
$for\ x \leftarrow 1\ to\ 8\ in\ steps\ of\ 1\ do$
$if\ bytes \leq h$
$bytes \leftarrow bytes + 1$
$continue\ next\ iteration\ in\ i\ loop;$
end if
$T = \sum_{l=(8-k)}^{8} (c_{[i][j][x]} \gg k)$
$S_{[i]} = S_{[i]} + T$
end for
end for
end for
end for
return T, the secret message
end procedure

8. EXPERIMENTAL RESULTS:

Using the above proposed algorithm, a sample hello world can be encoded in five different cover images as mentioned below in Table 1 and the same can be sent over the network to the receiver to complete the communication cycle.

Table 1. Sample text interpretation in the proposed algorithm using five cover images

Secret Text	ASCII	Binary Equivalent	Image-1	Image-2	Image-3	Image-4	Image-5(RS)
h	104	01101000	01	10	10	00	01
e	101	01100101	01	10	01	10	00
l	108	01101100	01	10	11	10	10
l	108	01101100	01	10	11	10	10
o	111	01101111	01	10	11	11	11
w	119	01110111	01	11	01	11	00
o	111	01101111	01	10	11	11	11
r	114	01110010	01	11	00	01	11
l	108	01101100	01	10	11	10	10
d	100	01100100	01	10	01	10	00

Let us assume that stego image-4 underwent an attack and the message digest mismatches with the content. On this contradiction, receiver will be able to identify that there was a corruption and recover the original text from other stego images as depicted below on Table 2.

Table 2. Secret text recovery on the loss of Image-3 due to steganalysis attack

Secret Text	**ASCII**	**Binary Equivalent**	**Image-1 (I_1)**	**Image-2 (I_2)**	**Image-3 (I_3)**	**Image-4 (I_4)**	**Image-5(RS)**	**Recovered Image-3 Ir = I1 ⊕ I2 ⊕ I3 ⊕ ... In ⊕)**
h	104	01101000	01	10	~~10~~	00	01	10
e	101	01100101	01	10	~~01~~	10	00	01
l	108	01101100	01	10	~~11~~	10	10	11
l	108	01101100	01	10	~~11~~	10	10	11
o	111	01101111	01	10	~~11~~	11	11	11
w	119	01110111	01	11	~~01~~	11	00	01
o	111	01101111	01	10	~~11~~	11	11	11
r	114	01110010	01	11	~~00~~	01	11	00
l	108	01101100	01	10	~~11~~	10	10	11
d	100	01100100	01	10	~~01~~	10	00	01

The comparison of results with data transfer without corruption/intruder's intervention are shown in Table-3 and Table-4 with their corresponding checksums.

Table 3. Comparison of checksum for the steganographed image, which transferred with corruption because of image color change

Image stage	**Checksum value(md5) (D)**	**Secret Text (ST)**
After Encoding	e64d69492b460cd25dbb42f970409f23	This is a secret text, which is hidden in an image file using steganography and having embedded checksum in it
After Decoding	e1c2e6f45c57978c86a78df764295972	Secret text got corrupted as the message digest are not identical

Table 4. Qualitative Comparison of proposed methodology

Parameters	**LSB methods (Existing)**	**Multi-image model Method (Proposed)**
Secret Text Recovery	No	Yes
Digest Inclusion on Stego files	No	Yes
Capability to identify MiM (Man in the Middle) attack	No	Yes
IPv4 Header Checksum check	Yes	Yes
Digest size used(md5)	0 bit	128 bits
Robustness	Less data loss	No Data loss
Integrity Check at receiving end	No	Yes

9. CONCLUSIONS

The proposed technique is effective on protecting secret message. Since the embedded checksum will validate for any unauthorized users or intruders corrupted the picture in any aspect. If any of the stego image underwent any steganalysis or MiM attack, then this proposed algorithm can regenerate the content of one stego image using other intact stego images received in the receivingend. Even if the attacker found the algorithm used for steganography in the stego picture by steganalysis and altered the quality of the hidden document, the tampering can be found by comparing the checksum if the same is obtained at the end of the recipient. Novelty of this approach, is the security of secret message is preserved and the model withstand a stego-image attack.

REFERENCES

[1] Petitcolas, FAP; Anderson RJ; Kuhn MG (1999). "Information Hiding: A survey" (PDF). Proceedings of the IEEE. 87 (7): 1062–78. CiteSeerX 10.1.1.333.9397. doi:10.1109/5.771065. Retrieved 2008-09-02.

[2] Dunbar, Bret. “A detailed look at Steganographic Techniques and their use in an Open-Systems Environment.” (2002).

[3] Y. Liu, S. Liu, Y. Wang, H. Zhao, S. Liu “Video steganography: a review” Neurocomputing, 335 (2019), pp. 238-250 (2019)

[4] J. Horng, C. Chang and G. Li, "Steganography Using Quotient Value Differencing and LSB Substitution for AMBTC Compressed Images," in IEEE Access, vol. 8, pp. 129347-129358, 2020, doi: 10.1109/ACCESS.2020.3009232.

[5] M. A. Hameed, M. Hassaballah, S. Aly and A. I. Awad, "An Adaptive Image Steganography Method Based on Histogram of Oriented Gradient and PVD-LSB Techniques," in IEEE Access, vol. 7, pp. 185189-185204, 2019, doi: 10.1109/ACCESS.2019.2960254.

[6] M. C. Kasapbaşi, "A New Chaotic Image Steganography Technique Based on Huffman Compression of Turkish Texts and Fractal Encryption with Post-Quantum Security," in IEEE Access, vol. 7, pp. 148495-148510, 2019, doi: 10.1109/ACCESS.2019.2946807.

[7] Setiadi De Rosal Ignatius Moses, “Payload Enhancement on Least Significant Bit Image Steganography Using Edge Area Dilation”, in Polish Academy of Sciences, Committee of Electronics and Telecommunication, Vol. 65, No. 2, pp. 287-292, 2019, DOI 10.24425/ijet.2019.126312

[8] Jagan Raj J and Prasath S. Article: Validating Data Integrity in Steganographed Images using Embedded Checksum Technique. IJCA Proceedings on National Conference on Research Issues in Image Analysis and Mining Intelligence NCRIIAMI 2015(1):5-8, June 2015

[9] J. R. Jayapandiyan, C. Kavitha and K. Sakthivel, "Enhanced Least Significant Bit Replacement Algorithm in Spatial Domain of Steganography Using Character Sequence Optimization," in IEEE Access, vol. 8, pp. 136537-136545, 2020, doi: 10.1109/ACCESS.2020.3009234.

[10] Jayapandiyan, Jagan Raj & C., Kavitha & Sakthivel, K. (2020). Optimal Secret Text Compression Technique for Steganographic Encoding by Dynamic Ranking Algorithm. Journal of Physics: Conference Series. 1427. 012005. 10.1088/1742-6596/1427/1/012005.

[11] Jayapandiyan, Jagan Raj & C., Kavitha, “Coalesced Technique in Steganographic Images Using Encoded Conversion for Augmented Security” in International Journal of Innovative Technology and Creative Engineering, Vol.7, No.4, April 2017

[12] Jayapandiyan, Jagan Raj & C., Kavitha, “Enhancing the data security and data integrity in steganographed images by store bit randomization” in in International Journal of Innovative Technology and Creative Engineering, Vol.5, No.12, Dec 2015

AUTHORS

JAGAN RAJ JAYAPANDIYAN: Jagan Raj is a research scholar pursuing Ph.D. in Department of Computer Science, Periyar University, Salem, Tamil Nadu, India. He had received his Master of Computer Application from Anna University, Chennai, India on 2009. Currently a member of ACM (Association of Computer Machineries). His major research interests are in Information security and steganography domains.

KAVITHA C: Dr Kavitha C is working as an Assistant Professor, Dept. of Computer Science, Thiruvalluvar Govt. Arts College, Rasipuram, Tamil Nadu, India. She has published 43 research papers on various national and international journals and her academic interest are in Image processing and Data mining research problems.

K SAKTHIVEL: Dr Sakthivel K is working as a Professor in Department of Computer Science and Engineering, K. S. Rangasamy College of Technology, KSR Kalvi Nagar, Tiruchengode, Namakkal, Tamil Nadu, India. His research areas are into Image processing and Data mining research problems.

PRODUCT RECOMMENDATION USING OBJECT DETECTION FROM VIDEO, BASED ON FACIAL EMOTIONS

Kshitiz Badola[1], Ajay Joshi[2] and Deepesh Sengar[3]

[1]Department of Computer Science, Mahavir Swami Institute of Technology, Guru Gobind Singh Indraprastha University, New Delhi, India
[2]Department of Electronics, Deen Dayal Upadhyaya College (University of Delhi), New Delhi, India
[3]Department of Computer Science, Riga Technical University, Riga, Latvia

ABSTRACT

In today's world, with the increasing demand of products and their growing productivity from producers, customers sometimes failed to decide whether they are interested in buying a particular product or not. So author, here proposed a framework which deals with the buying of only items of interest, for a consumer. In our feature-set, whenever any consumer tends to watch any video from YouTube, it results in breakdown into several frames (frames per second), and from there we use object detection technique to detect each and every object in a particular frame, and then to find whether our consumer is interested in that particular object or not, we use facial emotion detector to check whether our user is happy, surprised, neutral or any other emotion. After viewing those products which are present in a frame of a video. Merging only those items of interest which were tend to fall for consumer's positive choices (emotions), we then used Amazon online marketing technique to recommend products selected by our framework.

KEYWORDS

Convolutional Neural Networks, Facial Expressions, Object Detection, ImageAI, Selenium, Machine Learning.

1. INTRODUCTION

Decades back when machine learning was not introduced, it was hard to predict any human being's interest, mood and ability. But with the growth of artificial intelligence and machine learning applications we can predict most of the things of any user like expressions, gestures, etc. Machine learning allowed us to train our data as well as our system according to what we want to predict from a user, which leads in performing of several tasks that would have otherwise required considerable human efforts to very much extent. Authors, in this paper provides a framework which will use few machine learning techniques, which will result in product recommendation only and if our consumer who is willing to buy product is interested or not in that product, this type of product recommendation is needed because many times consumer fails to decide his/her interest in buying a product. In this framework we use video object analysis to find objects in a particular watching video, by consumer. The chosen video by the consumer from YouTube has been divided into several frames with the help of our code work, 1 second for one frame (e.g. for 2 minutes video, 120 frames), then these divided frames are then used by object detection model to predict all kind of objects/products. From there we pull out maximum

David C. Wyld et al. (Eds): SIGI, CSTY, AI - 2020
pp. 51-56, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.102006

occurring (counted) objects only. Then using another technique of CNN (Convolutional Neural Network), emotion detector model is used to predict consumer interest (like happy, sad, neutral, surprise, fear, disgust and angry) and finally with those interest factors, one or more products are targeted for prediction in recommendation from Amazon online store. Our purpose for choosing this project was not only to make consumers meeting with products of their positive choices but also, to ensure the increasing productivity in the market for only those products/materials for which consumer's interest is more severe/maximum, this will be highly useful for producers too. And this will also help in better economic growth and better substantial development over worldwide.

2. TARGETING MILESTONES

We used object detection technique to detect each and every product in the video using every frame per second from that video watching by user at that particular time period. We maintained a counting system to select maximum occurring object in a particular video. To ensure the maximum accuracy of product we manually removed human beings from detection criteria so that detection could be done on products only. Facial emotion detector used in this project is using CNN [1]to classify of data into different labels[2] [3] to provide user's emotion time to time and categorise them into seven categories like happy, sad, angry, disgust, fear, neutral and surprise. And from those we use only happy, surprise and neutral emotions to find the interest of any consumers watching that video as these three emotions were usually common for a positive interest showed by a consumer in any object at the time of purchasing/buying. The purpose for choosing facial emotion detector for this product recommendation project was because many previous research papers used audio detection [4], speech to text detection [5], object detection [6], etc. in a video to find an object of interest, but in this paper authors tried different strategy to recognise the current mood (emotion) of consumer to decide his/her product of interest. At last after having the results from emotion detector linked with object detector for a particular object (one or more than one product) is then used in recommendation using Amazon online store directly splashing those items of interest into Amazon page so that our consumer can now easily select his/her interested product for ordering/buying that product. Non interested products will not be selected by our framework because they showed negative choice of interest of a user like sad, angry, disgust and fear, so we ignore such objects/products and don't open Amazon store link for such cases.

3. OBJECT DETECTION

For accessing object detection model we need an image and to make that model work in our project we needed several images or frames because we are targeting videos from YouTube. So to get those frames, we used Selenium library to access the YouTube platform, then we set accurate pixels of a particular scene from that video(to focus only on the video part), leaving the background of YouTube behind, undetected and not selected to be a part of a frame. So that, now we are able to focus on our product with more clarity and accuracy. After selecting pixels for frames of a particular scene from that video we designed our code to capture frames continuously (per second) till the time our work is terminated and store those frames in other working directory of our project. What if our consumer tries several videos of many different products? Our work will get every frame from every (as many) video he/she is watching for either one or more than one product.

Now we used object detection for predicting different products. Model here used is a retrained model, i.e. resnet50_coco_best_v2.0.1.h5 [7] to save our training time, also we used FAST-RCNN (Region based Convolutional Neural Network) [8] system in our project, so that we could

predict our products from video frames. We used imageAI library which will predict all the objects in a selected frame then return the objects with its respective coordinates. Also to improve our object detection model we omitted/ignored human beings detection from the frame, just to focus only on products. After getting all the objects by their name and their respective coordinates from the frame we find out the area of each object in that frame and built our model in a way such that objects with larger areas in a particular frame should be selected because consumer's interest will most probably lean towards the object that are with huge area in a particular frame of the video. After selecting single or multiple objects from a frame as a result, we tried to find the number of times a particular object has been displayed in the video and with their count we proceeded for further recommendation process. For that authors maintained a list where objects from all frames are present with their respective counts, and highest number of count will be considered as prediction. For example, if we are watching two videos (one after other), an unboxing of laptop and P.C. from YouTube, object detection will result in targeting monitor, keyboard, C.P.U., etc. from that video. For more clarity, let us suppose any user watched two videos of different topics, like one video of microwave and other video of phone, the result after running of our framework will be of objects like oven, microwave, charger, phone, screen, etc.

4. FACIAL EMOTION DETECTOR

Our next milestone was to predict emotions of a consumer while watching any video and for that authors used TensorFlow to train the model. Kaggle dataset [9] was used in our model which was trained by several competitors at the time of competition in year 2013. Competition itself was conducted by Kaggle only, after giving an overview for our model, we found 3500 datasets in which we trained 2800 successfully and rest dataset were used as testing. Authors predicted accuracy was 92.10% on 25 epoch with the validation loss of 21.96% and with knowledge of Deep Neural Networks (DNN) we used dataset named Fer2013 and we design our own neural network which is depicted in figure 1. Authors at this point, named this model as 25.h5, now this

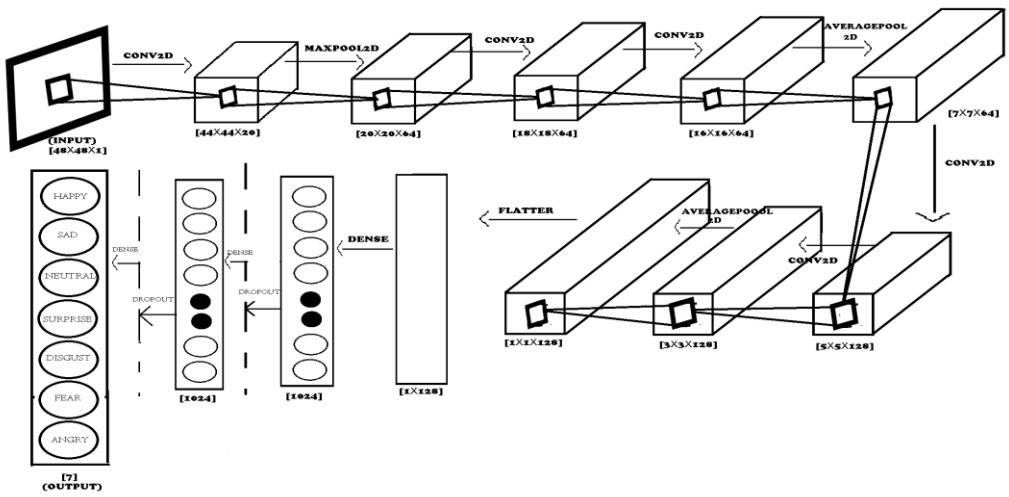

FIGURE 1. Model description

model will take input feeds and will predict output as classified between 7 categories that are happy, neutral, surprise, sad, angry, disgust and fear. Then according to our project we needed positive interest for a consumer so we mainly targeted in 3 emotions happy, neutral and surprise. Those 3 emotions will help in product recommendation as they are likely to be termed as

emotions for positive interest. If our user gets sad, angry, disgust or fear types of emotion while watching any video, it clearly results in his/ her negligible interest in that video so we will not provide product recommendation for such cases, this part is also important, because for good product recommendation system, all kinds of needs and emotion of a consumer should be judged properly and wisely. This model will show updated results time to time after executing it, till the time of termination so that our users need for each product could be refreshed at all time. And list of interested products could be maintained for product recommendation from Amazon online store.

5. MERGING RESULTS AND PRODUCT RECOMMENDATION

For merging results of selected product from frames and decided emotion, we used outputlinks.py. This will lead in reading of file and if our user is looking interested (happy, surprise or neutral) in any object from a frame of a watching video section, it will present a URL (Uniform Resource Locator) which will consist of a product that we obtained from object detection model based on our positive emotions (from emotion detector) linked with Amazon online store. So that anyone, anywhere in the world could access this product recommendation system for his/her interest directly with the product link available in Amazon store (if product not available at Amazon, then no recommendation). We are using selenium to access Amazon store from a chrome browser and from there our consumer can directly buy his/her product. What if we land in a position where we are interested in two or more products in a video? For that our work will present as many links, and open those links automatically on new tabs, and generate those URLs (as discussed) for only of products (one or more) that comes under consumer's positive interests according to our framework.

6. RESULTS AND DISCUSSIONS

Thus to ensure our working project based on user's individual product preferences, complex connections have been made and we surveyed 30 people (consumers) for our project and out of them we were getting positive results from 28 consumers for our product recommendation from Amazon store successfully only with the products of positive interest from a consumer. Our project for the recommendation of interested products using consumer choice is now successful. This framework if, used by any company/organization can lead in their better products advertisement as they will be targeting for consumer's product of interest. This framework can also be used in creating different types of advertisement content as they can trace the latest interest of the audience and their product preferences. For our working framework, we also present a block diagram of complete model which is depicted by figure 2.

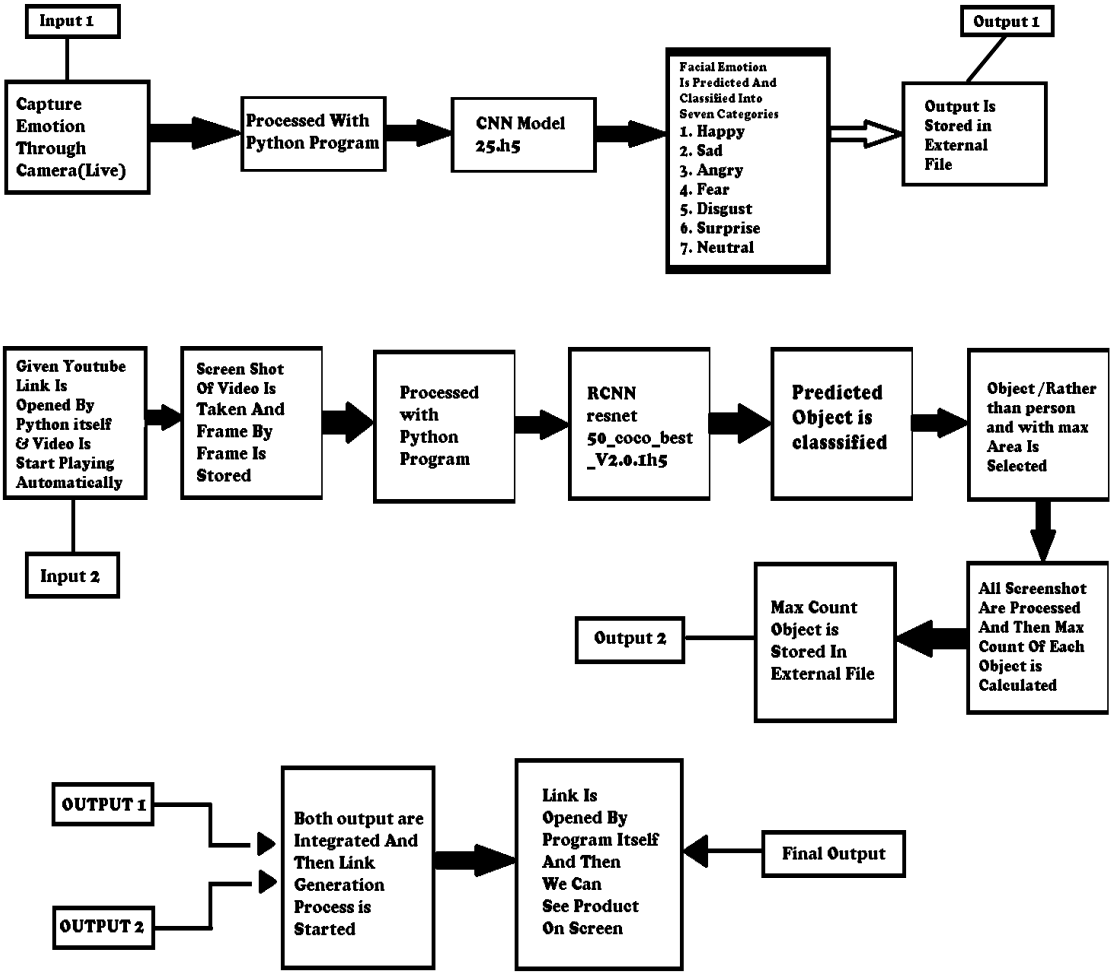

FIGURE 2. Final block diagram for working of project

7. CONCLUSIONS AND LIMITATIONS

As for now, our project is working successfully in recommending products of consumer's choice using their facial emotions and object detection, from any YouTube video. For the future work, we would like to combine all the feature-set and make it work for the real-time database system, not only YouTube video but also for other platforms like Facebook and Instagram posts. Also we would like to improve accuracy in our project by adding more classified emotions. We can make our own database for product recommendation where we can track each and every updated record of products of interest. Also we would to like to add more features in this for capturing reactions of consumers for products like audio detection, speech to text detection and any other machine learning technique which will lead in improvising our project accuracy

.

With our working project, there are some limitations that can be faced by any user at time of working on our framework, that are, our user who is watching video must be watching a product based content (for detection of product), otherwise no product will be selected for product recommendation for e.g. if any user watches fighting scene in a video, our framework will fail to detect any product from that scene. Another drawback could be, the API for our project is not

ready yet, so it will be difficult to run this work in any android system. There is one more minimal drawback which could affect our project, i.e. only a good processor P.C. would be highly appreciated for running our complete project, otherwise there is a chance of lagging if run in a slow processor system.

REFERENCES

[1] Yu, L., Li, B. and Jiao, B., (2019). Research and Implementation of CNN Based on TensorFlow. *IOP Conference Series: Materials Science and Engineering*, 490, p.042022. Available at: https://iopscience.iop.org/article/10.1088/1757-899X/490/4/042022.

[2] Ertam, F. and Aydin, G., (2017). Data classification with deep learning using Tensorflow. *2017 International Conference on Computer Science and Engineering (UBMK)*, Available at: https://ieeexplore.ieee.org/document/8093521
.

[3] Shiddieqy, H., Hariadi, F. and Adiono, T., (2017). Implementation of deep-learning based image classification on single board computer. *2017 International Symposium on Electronics and Smart Devices (ISESD)*, [online] Available at: https://ieeexplore.ieee.org/document/8253319.

[4] Smadi, T., Al Issa, H., Trad, E. and Smadi, K., (2015). Artificial Intelligence for Speech Recognition Based on Neural Networks. *Journal of Signal and Information Processing*, 06(02), pp.66-72. Available at: https://www.scirp.org/journal/paperinformation.aspx?paperid=55265.

[5] Nassif, A., Shahin, I., Attili, I., Azzeh, M. and Shaalan, K., (2019). Speech Recognition Using Deep Neural Networks: A Systematic Review. *IEEE Access*, 7, pp.19143-19165. Available at:https://ieeexplore.ieee.org/document/8632885/.

[6] Galvez, R., Bandala, A., Dadios, E., Vicerra, R. and Maningo, J., (2018). Object Detection Using Convolutional Neural Networks. *TENCON 2018 - 2018 IEEE Region 10 Conference*, Available at: https://ieeexplore.ieee.org/document/8650517.

[7] GitHub.2020. Release Models For Image Recognition And ObjectDetection • Olafenwamoses/Imageai. Available at: https://github.com/OlafenwaMoses/ImageAI/releases/tag/1.0/.

[8] Girshick, R., (2015). Fast R-CNN. *2015 IEEE International Conference on Computer Vision (ICCV)*, [online] Available at: https://ieeexplore.ieee.org/document/7410526.

[9] Verma, R., (2020). *Fer2013*. Kaggle.com. Available at: https://www.kaggle.com/deadskull7/fer2013.

AUTHORS

Mr. Kshitiz Badola pursuing B.Tech, Computer Science course from Guru Gobind Singh Indraprastha University, India. He had applied various A.I. tools in developing android systems, also he has immense knowledge in Deep Neural Network and Natural Language Processing. He is an assistant teacher of computer science department of his university.

Mr. Ajay Joshi graduated (by 2020) B.Sc. (hons) Electronics from Deen Dayal Upadhyaya College, University of Delhi, New-Delhi, India.

Mr. Deepesh Sengar is pursuing B.E. Computer Systems course from Riga Technical University, Latvia. He applied various programing tools and made the algorithm more efficient. He is the founder and lead developer of Quiqle Corporation and a tech enthusiast.

CLINICAL ASSESSMENT AND MANAGEMENT OF COVID-19 PATIENTS USING ARTIFICIAL INTELLIGENCE

Rashmi Phalnikar[1] and Subhal Dixit[2] and Harsha Talele[3]

[1]Associate Professor, SCET, MIT WPU, Pune, India
[2]Director ICU, Sanjeevan Hospital, Pune, India
[3]Assistant Professor, KCES's COEM, Jalgaon, India

ABSTRACT

The COVID-19 infection caused by Novel Corona Virus has been declared a pandemic and a public health emergency of international concern. Infections caused by Corona Virus have been previously recognized in people and is known to cause Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). Unlike the earlier infections, COVID-19 spreads alarminqly and the experience and volume of the scientific knowledge on the virus is small and lacks substantiation. To manage this crisis, Artificial intelligence (AI) promises to play a key role in understanding and addressing the COVID-19 crisis. It tends to be valuable to identify the infection, analyse it, treat it and also predict the stages of infection. Artificial intelligence algorithms can be applied to make diagnosis of COVID-19 and stepping up research and therapy. The paper explains a detailed flowchart of COVID-19 patient and discusses the use of AI at various stages. The preliminary contribution of the paper is in identifying the stages where the use of Artificial Intelligence and its allied fields can help in managing COVID-19 patient and paves a road for systematic research in future.

KEYWORDS

Artificial Intelligence, COVID-19

1. INTRODUCTION

In Wuhan, China a new corona virus was recognized in December 2019 and by 11th March 2020, a worldwide epidemic was declared by the World Health Organization (WHO). Corona virus is a group of infections and can take the form of SARS, MERS and COVID-19 that affects the respiratory tract. It has been scientifically proven that Corona Virus spreads via close contact and through respiratory droplets formed when a person sneezes or coughs. It may also be spread through fomite transmission such as infected surface and via physical contact with the body's mucous membranes like eyes, mouth, nose that might bring the pathogen into the body. The side effects of COVID-19 are vague, varying from asymptomatic to severe pneumonia and mortality. The most widely recognized clinical indications are fever and cough. Reverse transcription polymerase chain reaction (rRT-PCR) from a nasopharyngeal swab is the standard process of analysis. It can be also recognized from a grouping of chest CT scan, side effects and risk factors which show features of pneumonia [10].

Artificial intelligence (AI) programs the system to intelligent plays an important role in different domains like Natural Language Processing (NLP), robotics, biometrics, speech understanding,

David C. Wyld et al. (Eds): SIGI, CSTY, AI - 2020
pp. 57-66, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.102007

machine learning, deep learning, computer vision etc. It emphasises on reasoning action, computation and perception by use of artificial neurons that connect points in an artificial neural network. Artificial intelligence (AI) is changing our way of life, aiming to imitate human intelligence by a computer/machine in different issues. The three main considerations influencing AI algorithms are the accessibility of historical data, regular information and high computational force. The main goal of AI applications in medical field is to study relationships between precaution, treatment techniques and patient outcomes. AI technologies have been enhanced and useful in treatment protocol, drug development, diagnosis procedures, and personalised medicine.

The various roles played by AI during the pandemic are early admonition and alerts, investigation and perception of spreading patterns, prediction and identification of epidemic of diseases, study and examination of the virus, expectation of disease rate and contamination pattern, real-time disease monitoring worldwide, quick decision-making to spot the effective treatments and medication disclosure.

This paper discusses the role of AI technologies in the combat of COVID-19 infections, The primary contribution of this work is in framing a standard process followed by trained doctors to identify the COVID-19 patient and secondly to identify research areas in Artificial Intelligence. The rest of the paper is arranged as follows: the next section gives a brief background of the COVID-19 followed by literature survey, work and the conclusion.

2. Artificial Intelligence and its Importance

Artificial intelligence (AI) has contributed in all aspects of our lifestyles and its applications are changing our way of life by mimicking human intelligence to tackle everyday tasks. It permits us to mine knowledge from data by recognizing and understanding data patterns and later repeat them on new data for the required task.

AI has been applied effectively in different fields such as computer vision, medical imaging using deep learning, text mining, natural language processing, internet of things and many more. AI algorithms have been made use of in combating of COVID-19 on many fronts; from screening, identification, treating to drug improvement [13]. We give a brief description of it below, however the authors would like to point out that a systematic method of understanding of use of AI in COVID-19 infection management has not been given due attention. Most of the work has been random, leading to a haphazard overlapping work at different levels.

A brief description of the applications of AI in COVID-19 is given below:

1. Text mining and Natural Language Processing: Authors Du, S. Et.al.[3] discuss a hybrid AI model for COVID-19 infection rate forecasting is proposed with Improved Susceptible infected model. The natural language processing module, the long short-term memory networks are incorporate with ISI model to make hybrid AI model for COVID-19 forecast. NLP is utilized to extricate semantic characteristics from associated information for example citizen's avoidance awareness, outbreak control method of governments. These characteristics provide as inputs to LSTM deep learning model to modify the illness rate predictions of the SI model.

2. NLP, Text mining & network analysis are used to examine a Multilanguage Twitter dataset. In the paper titled "Understanding the perception of COVID-19 policies by mining a multilanguage Twitter dataset" by [4], they perceive the regular reactions to the pandemic and how these responses outbreak across time and countries. This dataset of 364K posts on

twitter have been extracted to give bits of knowledge on open reaction towards the pandemic over a few nations and various dialects.

3. Deep Learning for Medical Imaging: In paper "Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT" [14] Medical Imaging assumes a significant job in the identification of COVID-19 patients those most badly affected. The overall performance of radiologists in interpreting chest CT images for discrimination of COVID-19 from viral pneumonia. AI methods mainly deep learning have been utilized to process and examine medical imaging data to help doctors and radiologists to get better diagnosis performance. Medical images such as Computed Tomography (CT), Chest X-ray (CXR) and Ultrasound (US) have been broadly used for diagnosis, follow-up, prognosis and treatment in the clinical management of COVID-19.

4. Internet of Things: Many papers have discussed how IoT is ending up being exceptionally useful in the combat against COVID-19. This technology may be useful to the epidemiologists for looking through patient and furthermore in distinguishing the people interacting with the patients. Compliance of isolate by the patients can be guaranteed. The patients who violate the isolate can be traced. Also, this innovation can be useful in giving ease to the clinical staff by remote checking of in-home patients.

5. Computer Vision: Numerous AI-based computer vision camera systems are conveyed in China and over the world to check mob for COVID-19 side effects and screen individuals throughout lockdown. Infrared cameras are utilized in china to scan mob for high temperatures at train stations and air terminal. Sometimes they used a facial recognition system to identify the person with a high temperature and whether that individual iswearing a surgical mask.It is suggested that those cameras can test 200 people in minute and could apprehend the ones whose frame temperature exceeds 37.3°

6. Robotics: The job of robotics in healthcare relating to control of the spread of the novel corona virus. Also it helps to minimizeindividual to-individual contact and to guarantee sterilization, cleaning and services like isolate. This could bring about minimizing the life threat to therapeutic workers and doctors taking an lively function in the control of the COVID-19 deadly disease.

The utilization of AI and data sharing standardization protocols is useful for better worldwide comprehension and management of health during the COVID-19 pandemic. AI methods can also demonstrate its great effectiveness in supporting managers to make better decisions for virus containment when loads of urban wellbeing information is gathered by information sharing across and between smart cities [5].

3. LITERATURE SURVEY

Initial revelations of COVID-19 with AI were made by BlueDot, a Canadian company. BlueDot recognized the spread as well as anticipated the spread of infection to different urban areas. AI can be utilized as an early epidemic warning framework. Many tools have been developed and the next section describes a few of them briefly.

Artificial intelligence (AI) technologies support the ability of imaging tools and clinical specialists. AI image acquisition can help computerize the filtering strategy and reshape the work process with insignificant contact to patients provided that efficient protection to the imaging technicians is kept in mind. AI application also developed efficiently by precise depiction of infections in X-ray and CT images which facilitate resulting measurement. Analysis strategies

used COVID-19 analysis together with segmentation, diagnosis, image acquisition. The combination of AI with X-ray and CT are used in the frontline hospitals so as to represent the most recent advancement of medical imaging and radiology combating against COVID-19 [1].

The function of CR based IoT explicit for the clinical area is called Cognitive Internet of Medical Things (CIoMT). It is investigated to handle the worldwide test. Idea of CIoT is most appropriate to present pandemic as each individual is to be associated and observed through a very big network that needs efficient spectrum management. This tool is used for quick identification, active monitoring, tracking, better triage and control without spreading the infection to others [6].

An AI analytic tool is utilized to differentiate COVID-19 from different kinds of respiratory illness within seconds by patient's chest CT scan images in China. This new model supports great potential to improve early identification, relieve the pressure off frontline physicians, isolation and treatment [8].

FluSense is a contactless syndromic examination stage which is used to figure occasional influenza and other viral respiratory epidemic like SARS. It comprises of a novel edge-computing sensor framework, models and data preparing pipelines to follow coughs, crowd behaviours and to guess daily illness. It utilizes microphone and a thermal camera to constantly distinguish speech and cough sounds in a real-time manner [9].

A deep learning model COVID-Net is planned to identify the COVID-19 positive cases from chest X-rays. It also speed up the treatment for those who require it the foremost. COVID-Net uses deep convolutional neural network to analyze corona virus from chest radiography images. Open repository data contain COVID-19 and a variety of lung conditions taken as a trained data. The author's shows it is by no means a production-ready solution and they called on the scientific community to build up its advance in specific to improve sensitivity. [12]

As discussed here, as the worldwide health crisis deepens, the medical fraternity explores and understands new technologies to monitor and controls the spread of COVID-19 pandemic. Many methods of AI and its technology are suggested to understand the phases and spread of the virus, identifies high-risk patients, as well as controlling and segregating potential high risk persons.

4. COVID-19 PATIENT MANAGEMENT PROCESS

On the basis of our detailed study of published papers and study that have come up after the COVID-19 outbreak, it is alarming to see how AI application and its use for patient management is not standardized. There is random work that comes up and some standard understanding of the flow of events in COVID-19 patient management is missing. The various roles played by AI for COVID-19 treatments and management is crucial and can range from disease identification, study and examination of the virus, expectation of disease rate and contamination pattern, real-time disease monitoring worldwide, quick decision-making to spot the effective treatments and medication disclosure.

Contagious diseases are brought by pathogenic microorganisms like virus, fungi, bacteria, parasites. When a patient is infected, he may be symptomatic or asymptomatic. Corona virus is asymptomatic disease which is transmitted by Human-to-human interaction through fomites like emitted droplets of sneezing, coughing, speaking within few meters of distance.

Based on our intensive study, Figure 1 below shows the general procedure followed by a medical fraternity to recognize COVID-19 symptoms of a suspected patient and further management of the patient till he is cured.

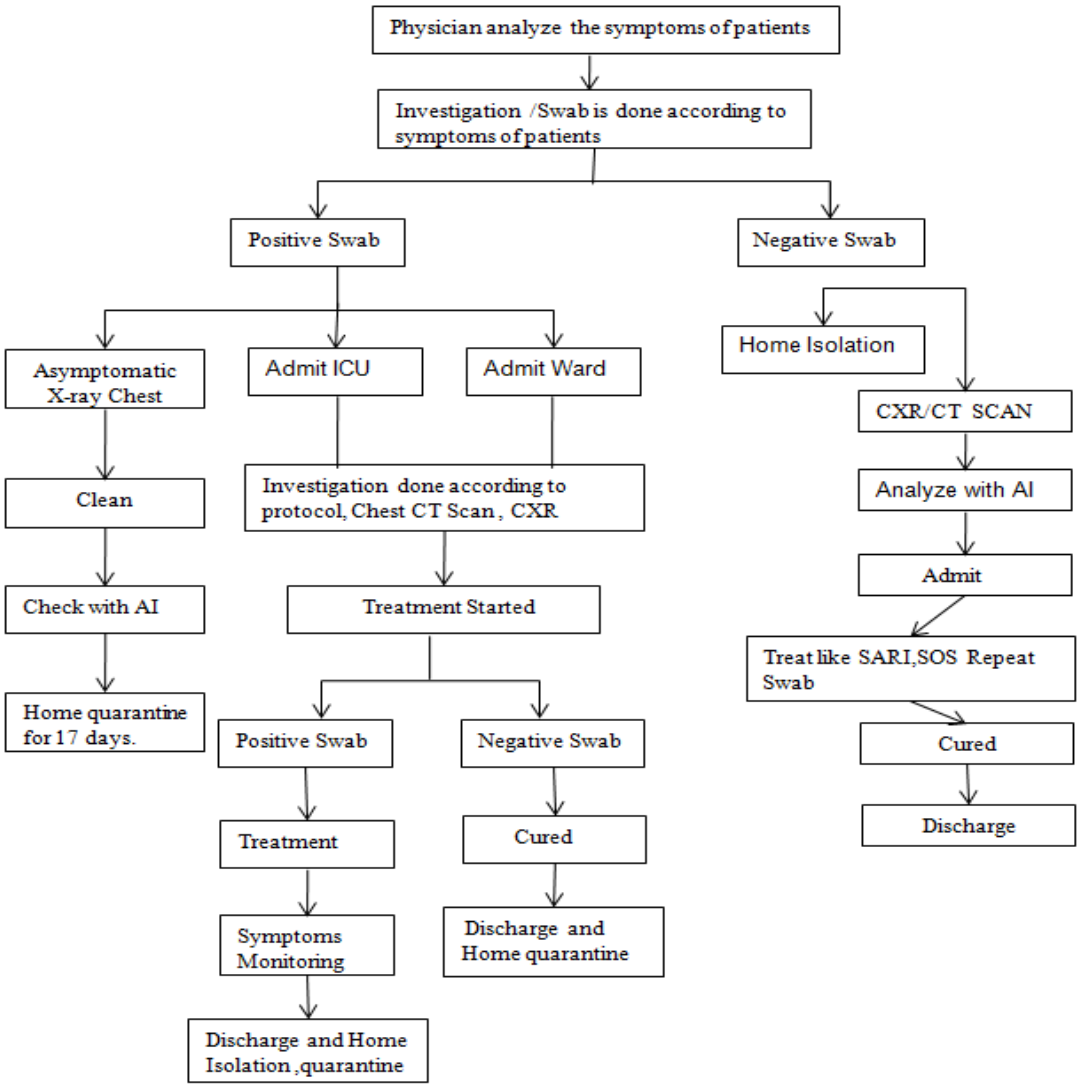

Figure 1. Process for COVID-19 Patient Management

Physician examine the symptoms of patients which include cold, fever, cough, dizziness, shortness of breath, new anosmia, diarrhoea, nausea or vomiting, headache, fatigue, muscle aches, throat pain and loss of smell and taste. These symptoms are also called are also called constitutional symptoms. Investigation of patients is done according to symptoms of patients and if these indications are strong, oral and nasal swab of patients is done. The symptoms may be categorized into mild, moderate, severe and critical.

Based on the symptoms and investigations, a inference can be drawn as to the severity of the infection. This analysis done after a detailed study and examination of hospital patients is summarized below in Table 1.

Table 1. Symptoms and Indicative Inferences of Covid-19 patients

Symptoms	**Clinical Severity**	**Indicative Inferences**
No clinical symptoms	Asymptomatic	No chest imaging findings
Mild clinical symptoms, such as fever, fatigue, cough, anorexia, malaise, muscle pain, sore throat, dyspnea, nasal congestion, headache	Mild	No abnormal chest imaging findings
Clinical symptoms as well as respiratory side effects of cough, shortness of breath	Moderate	Chest imaging showed mild pneumonia manifestation
Clinical symptoms and suspected respiratory infection symptoms plus any of the following: Shortness of breath, oxygen saturation	Severe	Chest imaging showed the lesions significantly progressed was a severe disease
Clinical symptoms and Cardiac injury, respiratory failure, hypertension, diabetes, cardiovascular disease, septic shock	Critical	Need mechanical ventilation and ICU monitoring treatment

Patients with mild illness may show side effects of an upper respiratory tract viral contamination. These comprise of mild fever, dry cough, sore throat, headache, nasal congestion, muscle pain. Patient with moderate sickness may show respiratory side effects of cough, shortness of breath. Severe pneumonia, acute respiratory distress syndrome present with severe symptoms of patients. A critical disease with features of cardiac injury, respiratory failure, hypertension, diabetes, cardiovascular disease may be developed in patient.

If along with these symptoms and conditions, the Swab tests are positive then admission to intensive care unit is required along with isolation or quarantine, depending on the moderate, severe or critical symptoms of patients. Under this isolation, the investigation of patients is done according to protocol. Patients can be feverish in the primary stages of infection with only respiratory and chills symptoms. In patients who are feverish without shortness of breath, a doctor would suggest estimations of complete blood count (CBC). CBC test includes haemoglobin, RBC count, WBC count. Chest CT is useful in recognizing viral pneumonia. When patient shows the symptoms of a fever, a patient needs to be taken for X-ray chest, CT and respiratory viral tests. Patients diagnosed with viral pneumonia require isolation and SARS-CoV-2 test. Further, if shortness of breath exists and oxygen immersion is seen, supplemental oxygen is prescribed and patient is admitted to an isolation ward. Special considerations are given for old people, immune compromised, pregnant patients. These patients should be treated as moderate or severe cases in the initial assessment.

If patient have any of the symptoms then test sample of nasopharyngeal specimen and Oropharyngeal specimen is taken by healthcare provider. Most tests use a cotton swab to get a sample from the inside of your nostrils and throat. If test is positive for SARS-CoV-2 then patient get treatments. After detecting positive case of patient treatment get started according to symptoms. Patients with COVID-19 require regular monitoring of vital signs like temperature pulse, respiration and SPO2. If swab is negative then patient get cured. Patient with negative swab are discharged from hospital. Patient provided with oral and written information guiding the need to stay in home quarantine.

If tests are negative then keep the patient in home isolation. Patient followed by mild symptoms may not require emergency interventions or admission to hospital if respiratory status is stable

and sanction discharge home. As per their clinical symptoms, patient management is done. Laboratory testing such as CXR, chest imaging ought to be performed in febrile patients. Chest computed tomography (CT) has been utilized to analyze COVID-19 showing that pneumonia is the most well-known appearance of the illness. So to detect the illness CT imaging is usually used.

We have described above the symptoms and indications. The goal of describing the entire process in details is to understand the future application area of Artificial Intelligence. We also expect that this flowchart will help further research work in understanding the methods used by medical fraternity to manage COVID 19 patient. During this entire process, large data may be collected at various stages and application of AI or Deep Learning on this data to discover patterns and make predictions can be very useful.

Artificial intelligence assisted with predicting which people would get the most sick and make sense of which of the side effects associate with severe acute respiratory infection(SARI). These patients have pneumonia-like symptoms. There might be a threat of contagion in these cases. So we chose to treat them like COVID 19 patients. So if required then repeat the swab of patient otherwise that patient get cured and discharged from hospital. But if these patients are tried positive for corona virus, they are moved to the COVID 19 ward and give treatment like COVID patients.

5. PATH TO FURTHER INVESTIGATIONS

Given below is a detailed analysis of the instances at which AI can play a role in further studies. Though much attempt has been made to use AI in COVID management, a systematic plan needs to be designed to understand the future of COVID analysis using AI, Deep Learning and advanced algorithms is missing. With reference to Fig. 1 as explained, the patients, who had positive tests, however do not show any signs are the ones who are asymptomatic at testing. As is logically obvious, asymptomatic people assume a major role in the transmission of SARS-CoV-2. Regardless of whether individuals have symptoms or not, they convey a similar measure of infection inside them. Symptom-based screening alone is unsuccessful to identify a high amount of contagious cases and insufficient to control transmission. In home quarantine AI is used to monitor home patients: patients can enter vital statistics like temperature, pulse, oxygen levels and this can be compared in a day to day analysis. If some alerts are detected then AI helps us to understand these alerts and the patient can be put under closer monitoring.

We can utilize Artificial Intelligence calculations to distinguish the infection using X-ray. If the patient’s X-ray is clean then patient is home quarantine for 17 days. The symptoms of such patients are check with wearable sensors.

Table 2. Analysis of COVID-19 using existing AI method

Sr. No.	Parameter	Analysis	Methods Used
1	Positive Swab	According to symptoms and history of patients admit to ICU/Ward	Decision tree, Random Forest, Neural Network, Classification and regression tree
2	Asymptomatic X-ray	Asymptomatic patients may have lung lesions on imaging so X-ray is preferred	XGBoost, Support Vector Machine
3	Investigation according to protocol	Symptoms and laboratory measurements were used to forecast the infection risk using Chest CT Scan, CXR	To predict the infection at early stage Support Vector Machine and back propagation neural network can be used. Can help in building effective predictive models to find various infections as early as possible.
4	Analyze with AI	Check respiratory status using CXR, CT scan	Support Vector Machine
5	Asymptomatic check with AI	To track heart rate variability, blood pulse wave, heart rate, blood oxygen saturation, respiration rate, activity, barometric pressure, skin temperature wearable sensors are essential to monitor asymptomatic patients.	IOT
6	Discharge & Home Isolation	Artificial Intelligence is used for remote monitoring of home-isolated patients through smart bracelets or smart phones. If the patient breaks quarantine, automatic alarm will sound and give a warning message.	Artificial Neural Networks can be trained to conclude qualitative characteristics that are depend on network inferences and intensity, which can be associated with the patient's condition

Other important methods that can prove beneficial are:

- AI has the influence to improve chest imaging beyond just screening for signs of COVID-19 in a patient's lungs.
- AI can help to improve risk classification of patients for the type of care they receive based on the predicted course of their COVID-19 infection.
- The COVID-19 database made up of medical data of real patients contains data of symptoms such as chest pain, muscle pain, chills, colds, discomfort, conjunctivitis, cough, diarrhoea, dry cough, dyspnoea, respiratory symptoms, weakness, emesis, expectoration, eye irritation, pneumonia, fatigue, dizziness, gasp, lesions on chest radiographs, little sputum, malaise and so forth. This can be useful for future predictions.

- Supervised learning algorithm such as Decision tree, random forest, neural network and support vector machine algorithms can be applied to diagnose the patient who is infected. Whether a patient is suffering from a disease or not based on symptoms is done by decision tree. SVM examine data and recognize a pattern which is used for classification and regression analysis.

- Order of treatment is important plays an important part of the procedure even on account of developing irresistible disease where endeavours must be prioritized.

- An AI system can be use where the patients having mild symptoms may be home quarantined and the patient is at higher risk, an appropriate drug is administered. A system can propose an antibiotic treatment to the patient depend on factors like the infection sites, body temperature, antibacterial spectrum, symptoms and complications. In the final step, based on the result, treatment will be given to patient.

In all the role of AI in a COVID-19 patient treatment must be given due thought and consideration.

6. CONCLUSIONS

Artificial Intelligence is powerful tool to recognize early infections due to corona virus. AI has contributed to managing the corona virus infection (COVID-19). It makes a difference in observing the condition of the infected patients. It is also helpful to facilitate the research on this disease via examining the existing data. It can make decision making and therapeutic plan by creating valuable algorithms. AI can offer assistance in creating appropriate therapeutic regime, precaution procedures and medicine and immunization innovation.

REFERENCES

[1] Feng Shi, Jun Wang, Jun Shi, Ziyan Wu, Qian Wang, Zhenyu Tang, Kelei He, Yinghuan Shi, Dinggang Shen (2020) "Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation and Diagnosis for COVID-19", *IEEE Reviews in Biomedical Engineering*, pp.1-13, doi: 10.1109/RBME.2020.2987975.

[2] Zhavoronkov, A., Aladinskiy, V., Zhebrak, A., Zagribelnyy, B., Terentiev, V., Bezrukov, D. S. and Yan, Y. (Feb 2020) "Potential COVID-2019 3C-like Protease Inhibitors Designed Using Generative Deep Learning Approaches", doi: 10.26434/CHEMRXIV.11829102.V1.

[3] Du, S., Wang, J., Zhang, H., Cui, W., Kang, Z., Yang, T., and Yuan, Q.(2020) "Predicting COVID-19 Using Hybrid AI Model"Available at http://dx.doi.org/10.2139/ssrn.3555202.

[4] Lopez, C. E., Vasu, M., and Gallemore, C. (2020) "Understanding the perception of COVID-19 policies by mining a multilanguage Twitter dataset", arXiv preprint arXiv:2003.13907 .

[5] Allam, Z., and Jones, D. S. (2020, March) "On the coronavirus (COVID-19) outbreak and the smart city network: universal data sharing standards coupled with artificial intelligence (AI) to benefit urban health monitoring and management"*InHealthcare*, Volume 8, No. 1, p. 46, MDPI.

[6] Swati Swayamsiddha, Chandana Mohanty (September-October 2020) "Application of cognitive Internet of Medical Things for COVID-19 pandemic", *Diabetes Metab Syndr*, Volume 14, Issue 5.

[7] https://www.infoq.com/news/2020/03/deepmind-covid-19/

[8] Chen, J., Wu, L., Zhang, J., Zhang, L., Gong, D., Zhao, Y., Hu, S., Wang, Y., Hu, X., Zheng, B., Zhang, K., Wu, H., Dong, Z., Xu, Y., Zhu, Y., Chen, X., Yu, L., and Yu, H.(2020) "Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: A prospective study", *medRxiv*, https://www.medrxiv.org/content/early/2020/03/01/2020.02.25.20021568.

[9] Hossain FA, Lover AA, Corey GA, Reich NC, Rahman T.(March 18, 2020) "FluSense: A Contactless Syndromic Surveillance Platform for Influenza-Like Illness in Hospital Waiting Areas",*Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Volume 4 No. 1, pp. 1-28, https://dl.acm.org/doi/abs/10.1145/3381014.

[10] Fatima M. Salman, Samy S. Abu-Naser, Eman Alajrami, Bassem S. Abu-Nasser, Belal A. M. Ashqar(March 2020) "COVID-19 Detection using Artificial Intelligence", *International Journal of Academic Engineering Research (IJAER)*, ISSN: 2643-9085 Volume 4, Issue 3, pp. 18-25.

[11] Niiler E. "An AI Epidemiologist Sent the First Warnings of the Coronavirus" WIRED [Internet].https://www.wired.com/story/ai-epidemiologist-wuhan-public-health-warnings/ Accessed April 1, 2020.

[12] Wang L, Wong A.(2020) "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images", arxiv:2003.09871.

[13] Chun A. "Coronavirus: China's investment in AI is paying off in a big way" [Internet]. South China Morning Post; 2020. https://www.scmp.com/comment/opinion/article/3075553/time-coronavirus-chinas-investment-ai-paying-big-way. Accessed April 4, 2020.

[14] Bai HX, Hsieh B, Xiong Z, et al.(2020) "Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT", *Radiology*, https://doi.org/10.1148/radiol.2020200823.

[15] Zeashan Hameed Khan, Afifa Siddique, Chang Won Lee (28 May 2020) "Robotics Utilization for Healthcare Digitization in Global COVID-19 Management", *International Journal of Environmental Research and Public Health*, doi:10.3390/ijerph17113819.

[16] Singh RP,Javaid M,Haleem A, Suman R (2020). "Internet of things (IoT) applications to fight against COVID-19 pandemic", Diabetes Metab Syndr., Volume 14, Issue 4, pp. 521-524, doi:10.1016/j.dsx.2020.04.041.

APPLICATION OF AUGMENTED REALITY & VIRTUAL REALITY IN ARCHITECTURE AND PLANNING: AN OVERVIEW

Ms. Pearl Jishtu[1] and Madhura A Yadav[2]

[1]School of Architecture & Design, Manipal University Jaipur
Dehmi Kalan, Off Jaipur-Ajmer Expressway, Jaipur - 303 007,
Rajasthan, India
[2]Manipal University Jaipur
Dehmi Kalan, Off Jaipur-Ajmer Expressway, Jaipur – 303 007,
Rajasthan, India

ABSTRACT

AR and VR – simulation tools created to assist global evolution for saving time. Time as resource is difficult to harness; however, it would make work highly efficient and productive when tackled with automation. All concerned are excited about AR and VR's involvement in our lifestyle, but not all have comprehended its impact. AR and VR in Architecture & Planning were introduced as assisting tools and has helped generate multiple design options, expanded possibilities of visualization, and provided us with more enhanced, detailed, and specific experience in real-time; enabling us to visualize the result of work at hand well before the commencement of the project. These tools are further being developed for city development decisions, helping citizens interact with local authorities, access public services, and plan their commute. After reviewing multiple research papers on AI, it was observed that all are moving forward with the changes brought by it, without entirely understanding its role. This paper provides an overview of the application of AR & VR in architecture and planning.

KEYWORDS

Virtual reality (VR), Augmented reality (AR), Architecture, Urban Design, Human Computer Interface (HCI)

1. INTRODUCTION

Architecture and urban design rely on a similar process that includes data collection and analysis, design decision, visualization, and evaluation. AR has been developed for immersive and collaborative use in all of these stages to analyze the best design solution outcome in a better manner. Some of these processes are already in the application. Apart from using the simulation for the professionals, the application of these systems is also being evolved for the use of the public, students in the field of design, and policy makers. A study of various research papers was conducted and summarized to review and analyze the current developments in simulated or algorithmic tools. The papers cover diverse topics under theory, concept, systems, tools, applications, and implications.

David C. Wyld et al. (Eds): SIGI, CSTY, AI - 2020
pp. 67-77, 2020. CS & IT - CSCP 2020 DOI: 10.5121/csit.2020.102008

2. METHODOLOGY

The research paper is a summarization report of utilization of AR (Augmented Reality) and VR (Virtual Reality) in the field of architecture and urban design, for an in-depth understanding of the potentials and challenges of the technology. The main objective of the research is to record the progress of these tools with a perspective specific to the design field and to understand its diverse applications. The information was compiled by going through 20 research papers on the subject cited, of various scholars, published in conferences and readily available on the research websites. The keywords were used to search for the subject related papers and were then studied in depth for a better understanding of the same.

3. LITERATURE STUDY

(Seichter et al. 2005) [5] Elaborated on understanding the association between human beings, the natural world and the successive insinuation on their interaction. This was related to Phenomenology, which is also related to the realm of Human-Computer Interfaces (HCI) and urban design. Phenomenology is the principle of empirical observation on events that are based on different theories but cannot be directly explained through them. Various theories have been put forth about the connection between perceivably real and virtual objects. Husserl and Gibson (1931) [1], Schutz and Northwestern University (1967) [2] and Norman (1988) [3] came up with theories elaborating the intangible connections between real and virtual objects, to establish a conscious sense of presence.(i) *Human Computer Interfaces (HCI),* of which AR and VR are the components, having a great scope in the field of urban design and architecture (Frazer et al. 1980)[4]. This allows the user to work in both real and virtual environments simultaneously within all dimensions, instead of the traditional physical models. Their inherent physical nature makes them easy to understand by both the layman and experts alike. However, they find limitations in their flexibility towards a dynamic presentation. Augmented Reality (AR) however, offers a variety of opportunities where urban design can gain from in the intersection of real and virtual realms. (ii) *Collaboration* is important in urban designing, and cognition being another feature is invoked through the utilization of immersive technology and perceptible interlinking. Sensations triggered by ‘grasping’ and ‘moving’ aid the designers’ cognitive process and their awareness of ‘being inside’, which in AR, is a mixture of real and virtual(iii) *Cognition* is the ability to comprehend and create comparison between data through analysis. Thus the understanding of the design and its relationship within the urban context is enhanced, site-specific contexts are better recognized, and various options can quickly be investigated. Since AR permits a blending of virtual entities with a real environment, it can, as a result, increase the comprehension of the design proposal. (iv) *Innovation*- Students and teachers can explore various theoretical and practical frameworks to understand dynamically and spatially complex relationships.

One of the study goals was to identify how designers create and communicate initial design ideas by employing AR. Using an AR system, designers gain a more complex understanding of their design relationships and engage in a richer communication with their partners. AR contributes to urban design in an innovative approach, thus enabling new forms of design expressions.

(Sørensen, 2006) [6] States that Augmented Reality is a mediating tool to design ideas between the design and a person. It elaborates upon the components of the AR system in principles, which include; (i) Position registration system to provide the exact geo-location for the placement of the 3D model. For interiors, infrared, ultrasound systems, or graphical markers system are explored with limited installation time. While the use of the exterior of differential GPS is accurate, the system is currently expensive, (ii) Orientation registration system – This requires the user's exact location and the user's angle. Currently this has limitations of a slow data processing, resulting in

lagging and inaccuracies in reading data leading to disproportionate models in the surrounding, and (iii) Display device – there are multiple viewing options, one of them being a headset, which currently has two different systems of operating. Ones consist of a transparent screen with the model visual on it and in the other the surrounding images are fed to the AR system without a transparent screen.

The limitations analysed of AR Systems in Architecture and Urban Design were, minor reception of the system's markers leading to poor visualisation due to the sun glare and in case of dynamic AR systems, the accurate positioning by the system leads to errors like disproportionate models. AR will induce new insights in architectural learning, introducing the students to new parameters that will enable them to immerse in their designs virtually. This will enable them to study alternatives at different abstraction levels, thus contributing to the earlier design phases better. Dynamic models and collaborative systems will broaden the scope, and it will be possible to change proposed structures or models in AR, and one can experience the simulation of changes in physical conditions, animate, and visualize pollution, noise, etc.

(Döllner et al. 2006) [7] Virtual city 3D models are integerated comprehensive systems utilizing diverse urban space information. The application of this standardised model is discussed in the perspective of various feilds that require visual representation of urban environment. The model is a virtual representation of terrain, building and vegetation models alongwith roads/transportation systems. It aims to incorporate three categories of data (i) Cadastral data delivers the footprints of buildings and land ownership along with the ownership status, (ii) Digital terrain model and photovisualisation, (iii) 3D Building model geometry captured by laser-scanning and photogrammetry based methods, and (iv) architectural models with their elements and historic references. There is no widely accepted standardized system but some of the systems in use include: CityGML, 3D Studio Max object files and VRML files, ESRI Shapefiles. The system is an important collaborative tool and is a great support system for 3D decision making. It covers functionality like managing, integrating, and distributing complex geoinformation supported a consistent communication metaphor, the virtual 3D city model. These functions are then divided up in sub-systems for content authoring, editing, storing, and presentation extendible 3D city model system. In addition, techniques for the automated mapping of 2D landscape plans and architectural plans to 3D geovirtual environments based on a heuristic-algorithm approach are investigated.

(Seichter, 2007) [8] Explored the aspects of co-located collaborative design process with the utiliztion of 3D models and AR tools. Experiments conducted to record the user response of these tools in the field of design communication, physical engagement and the concept of presense. Factors of tangibility and directness of the user input device were used to distinguish the two experimental conditions. The observations suggested that the involvement of new techniques and tools for a better understanding of the inter-spatial relation are beneficial for design, with a better and an effective output. The adoption of novel methods like the TUI keeps on linking the bridges between reality and virtuality. The investigation has raised scopes for the future development to incorporate these technical interfaces to achieve a collaborative settting with a consideration for presence and human perspective.

(Wang, 2009) [9] Discusses the relation between MR (Mixed Reality), VR (Virtual Reality) and AR (Augmented Reality), while futher elaborating upon their potenials, applications and challenges in the field of architecture and design. Some of the software and hardwares were disscused with their current application for the various stages of design were Sketchhand+, Benchworks, ARTHUR, AR CAD (Augmented Reality Computer Aided Design), MRCVE (Mixed Reality-based Collaborative Virtual Environment), Studierstube, BUILD IT and TINMITH2. Even though there is plenty of research and development in the field of AR and VR

simulations, but for specific application like architecture and urban design the field is yet to be explored in detail. The main reasons of this stagnation are (i) lack of industrial domain knowledge, due to the inavailibility of a well organised integerated 3D database providing information of the mechanical services, (ii) creating a well defined reality model, and (iii) the accuracy and the robustness of the current tracking systems.In brief AR system is not very flexible and specific in current use and needs a pre analysis of the environment. The future robust AR aims to have a system that could work in unstructured indoor and outdoor environments.

(Abboud, 2014) [10] AR and VR has a wide range of toolsets that help generate simulations, out of these subsets the mobile augmented reality (MAR) is one option which is widely explored in practical research. The MAR system has diverse applications across design, construction and post construction that provides us with in-depth data. These include (i) interaction of the clients, consultants and designers at the initial design stage, (ii) the BIM data can be directly overlaid at the site, to oversee any later stage issues and it could be efficient in locating materials, equipment and safety zones. It could assist the labours to carry out complex work by aiding them with in-situ visual assistance. On site model work progress could also be looked over and reported to the site, preventing causalities, and (iii) Post-Completion the facilities to maintain the building structure can take aid of the augmented system to locate the services and facilitate the maintenance requirements. These MAR applications face technological, social and financial issues which need to be resolved before moving forward. Tracking systems need to be evolved and specific for locating services, specialists will have to train the workers with these systems and new protocols will have to be implemented for maximum efficiency.

(Seichter, 2020) [11] Explores the visualisation of urban data beyond the current 2D use, incorporating the tangible and non-tangible features which are essential to understand the human interaction in an urban space. The interaction techniques and technologies incorporate the significant contribution of BenchWorks that is collaborative and immersive AR-enabled working environment. BenchWorks is developing the tools and technologies to provide an immersive AR experience specifically for Urban Design data collection and analysis. Their *Benchmark sketchbook+* has turned out to be a very useful tool for the initial design decision making process. ARToolKit [12] provides an easy way to create input devices that do not rely on wires and therefore are more convenient in a multi-user setup. The magnet tracking provides higher precision for head orientation and, consequently, provides a more stable augmented vision. Major benefits for this early prototype are combined: the precision of a magnetic tracking system and the freedom of tangible interfaces overcome occlusion within the scene.

Shared Tangible City - Choice and combination of media play a key role in the investigation of space and, subsequently, architecture. BenchWorks goes beyond that, creating a toolkit for investigating a site's spatial feature dynamically in a setting with local and remote collaboration capabilities. The way a user can design in BenchWorks is by representation of void and non-void space of the city, adding volumes and making notes; the interaction being tight-knit with tangible interfaces. The toolkit uses two different tools: toolbox and the modifiers represented by pens and the sponge.

(Kim et al. 2011) [13] used an AR system in a studio setting to introduce shared learning as an essential part in urban design, where students are conversant in syntax and grammar. AR is employed to dynamically and spatially understand complex relationships, which helps people understand urban issues during a new and creative way. The experiment conducted helped to understand the participants' sense of presence and their level of immersion. AR conceptualisation gave an in-depth understanding of the algorithmic framework and modelling that are adopted and the various challenges that are to be faced just like the sun glare (Chung et al. 2009) [14], temporary lags created in the systems (Wang and Dunston 2005) [15], heavy equipment (Santos et

al. 2007; Chung et al. 2009) [16][14] and therefore the human factor. These concepts and ideas implemented with hardware or software tools that have their own challenges like(i) the human interaction is not very experiential, (ii) potential use of agent-based is yet to be explored (Belcher and Johnson 2008)[17], and (iii) portability of the entire system is still a substantial limitation. While evaluating the whole system for the quality and the productivity through quantitative survey, cognitive and effective workability were the common issues noted. Even though a lot of literature survey and practical study was done on this subject, a wide range industrial adoption was still found lacking. This was mainly due to the expensive and heavy toolkits that were not designed specifically for the design field.

(Schubert et al. 2015) [18] Design Communication in the form of 2D drawings and details is only comprehendible by the professionals in the field, but not the clients or the laymen constructing the structure. Design decision making at the initial stages becomes easier with illustration drawings which are currently used in the digital format. Now, with the intervention of AR and VR systems it would bring better results. The idea would be to simulate the master plan or the building geometry on a table top VR or could also be augmented on the site via headset. The two design components discussed are the collaborative design platform and the on-site AR-applications. The second is discussed on the basis of its potential to grow and evolve, as it is bound to make certain processes of design better.

(Thabet, 2002, updated 2015) [19] The foremost important benefit of VR is that, the clients can rehearse a building's design experiencing the finished product as if they were in the actual space. Flaws within the designs can also be detected easily with VR. Certain concerns raised were, (i) use of 3D and VR are often beneficial as long as used early at the conceptual design process, (ii) housing developers did not want to expose users to views from sensitive unrealistic angles, and (iii) when dealing with local authorities to avoid exposing the design or construction to other issues where an advanced visualization must be used with care. The workers can utilize AR to detect hazardous faults and prevent causalities. Retik and VIRCON are two automated simulations that deal within the constructional perspective. The complete analysis of the structure just in case of calamities could be conducted with the assessment of the relief shelter within the particular building. The utilisation of these tools is to simplify the upcoming complex designs.

(Milovanovic, 2017) [20] Highlighted two aspects of AR and VR design i.e., interaction and immersion. The article in the initial portion discusses the analysis derived from research papers regarding the development of AR and VR in the various design processes and its subsequent implementation. It was observed that maximum research papers were about the development of the hardware/software development of the system, and then followed by design; however, the applications are very few to date. It helps us analyze and understand that due to certain limitations like the cost of technology, we are stuck at the conceptual and design stage. The five significant systems elaborated upon this text include; (i) HDM AR- as proposed by the Benchmark system, (ii) HYVE 3D - a system that is an interplay of Computer-Aided Design (CAD) and Virtual Reality, resulting in the formation of VRAD (Virtual Reality Aided Design), making collaborated design in the initial stages feasible, (iii) HMD VR, used in the CAP VR system -The environment proved to enhance students' design quality. Students' immersion in the virtual model of their design augmented their spatial comprehension and enriched their evaluation. However, the communication with the tutors was not seamless because students were wearing an HMD, (iv) Tangible AR – the system illustrated by the Luminous Planning Table, and (v) SDAR or SAR – It utilizes the joint surfaces available to us around like the walls and floors to be used as display devices, providing the design on a 1:1 scale experience.

(Broschart, 2015) [21] Studies the varied tools that are being used or being developed for the use of urban planning and architectural visualization. Keeping in mind the various limitations that are

foreseen with the use of AR systems (portability of computers and time required for installation), this research lays focus on the smart portable i.e., the mobile phone which consists of four-element components (1) Computer software generating the 3D model that is to visualized, (2) Geo locating and extracting the exact location data, (3) detecting the user's orientation and (4) the last would be overlay these and finally display them. Its diverse applications include: (i) *Campus Navigation* – Layar tool is utilized to navigate campus with geo-location satellites, (ii) *IGA 2017* Berlin – Layar was used to create a digital walk throughout the built area, albeit the system came with barriers. Therefore, RADAR-system (Resource Annotation and Delivery for Mobile Augmented Reality Services) was introduced to breach the barriers and permit the users to place their own location data. Even these apps were limited by the large streaming files and the availability/accessibility of the internet, (iii) *Building Culture Saarbrunkrn* – to overcome the limitations of the GPS based AR systems, alternative systems based on image recognition tracking systems are also used. "Layar" published a version known as "Layar vision", with its first application being tested in Saarbrunkn. The system was updated with the city street photos of 1950s, with colour marking the architectural elements that made the city legible and were in the pre-set time covered with hoardings or had been destroyed. The experiment was done at various points throughout the city to involve the local people and letting them experience the lost heritage and what the city aimed to get back, and (iv) Augmented Plan – when it involves the mediation of the plans that we present to the client, most of the information we lay in front is tough to comprehend. As a solution to this, apps can take the virtual data such as plan and augment it within the client's phone itself.

(Redondo et al. 2011) [22] VR and AR research have shown great potential in the field of design education. However, not much progress has been made in the field of education through these tools, due to technological issues and device costs. In Architectural and urban design the understanding of spaces is an essential component. Currently, this knowledge is achieved during the latter stages of a professional's carrier, while with the help of AR the students could be introduced to these scenarios during the initial part of the course itself. It will result in better understanding for the students compared to the traditional tools used to study architecture that are still majorly limited to 2D. It is essential to explain the work process in architecture to juggle between multiple other disciplines to make a project feasible. Traditional representations embody both a need for abstraction and interpretation and can thereby be challenging to understand. Therefore, they possibly mislead those not trained in the field. In response to this, architectural projects and urban design drawings are complemented with perspectives, text and 3D models, both physical and digital. Still, there is a common problem of understanding both spatial qualities and the scale of proposed buildings and structures. As a result, project-owners, politicians, decision-makers, and the public often have expectations based on their interpretation that differ from or exceed the realised project's qualities in its context. In urban design, it becomes vital to have a human-spatial experience. Changes we need to look forward to (i) prior to realisation, complete 3D models of proposed projects can be studied regarding all aspects. Today we build 3d models, after AR we will be able to experience them and in future maybe make in situ changes, (ii) one can experience simulation of physical conditions changes, animate and visualize pollution, noise, etc., (iii) human factor studies will be made tangible and better outcomes for the design will be achieved, and (iv) public opinion and suggestions could be incorporated as through these tools, design can be accessible to all.

(Van Dessel, 2018-2019) [23] Virtual Reality deals with a holistically created or designed models; it is well utilized with the augmented reality which uses real time location. Traditionally in urban design cardboard is used as a means to create physical models, which is not a very flexible option and the visualisation of the space is done only from a bird-eye view. City Model – creating a virtual model to immerse in the design and make changes along the visualisation is one of the

main aims of the VR. The main advantage is to be able to see the proportions and height which otherwise cannot be visualised in the 2D GIS system, currently in use.

Three visualisation techniques have been adopted (1) – Photorealistic Visualisation aims at providing a real like impression of the environment, mainly for the non-experts. The complex geometrical forms of the environment e.g., trees increase the file size and create multiple issues. (2) – Visualisation of Information and Data which enables information to be comprehended and immersed with the city model e.g., occupancy, year of construction, etc. (3) – Illustrative and non-photorealistic visualisation can be used for creating the first look of the model, providing with the urban information, helping in decision making process.

City GML – it is a standard core model that acts as a source of information for multiple disciplines which require urban data. Although the standardised model does provide a clear structure, the coding languages are unknown to the urban planners. This becomes a little cumbersome and restrictive to the designers.

(Ahmed et al. 2019) [24] The recent development in information technology has huge opportunities to improve the architectural education in terms of methodologies, strategies and tools. Studies explores the importance of using virtual reality with open minded and progressive methods in teaching architectural courses and not just stick to the traditional methods of yesteryears. The research aimed at evaluating the (BC\VR Software) in architectural education of building construction courses as a case study at Jordan University of Science and Technology (JUST). This Building Construction using Virtual Reality technology (BC\VR) is computer software designed by the authors for research purpose and presents 4D models (3D model and time dimension) for certain building construction phases using VR technology to do immersive and non-immersive virtual reality experience for the users in terms of three axes: (i) providing students with the building construction information, (ii) achieving enjoyment, and (iii) integrating with other courses. The data was obtained from architecture students on a structured questionnaire. The results were highly in favour of VR software than those of the traditional teaching methods. As conclusion, the use of BC\VR software as tool in building construction courses is very useful and effective for the students. The VR technology can also find application in many other courses offered in the curriculum of Architectural programs. VR integrates the traditional way of education by bringing the real world into the classroom, however, its application in architectural educational purposes have been limited so far.

(Gębczyńska-Janowicz, 2020) [25] The evaluation of research presented in this article indicated its effectiveness as an educational tool in the architectural design studio to improve learning. Realistic concept visualization enables students to experience the special consequences of proposed solutions. VR is able to support the design of commemorative places by applying reconstructed buildings to the real image. With virtual reality, the guide could use tools to create an augmented reality experience, being able to draw attention to certain elements that otherwise may not be easy to explain in real life. CAVE is one such interesting tool for an immersive design environment. Virtual Reality (VR) technology allows architectural design classes based on advanced digital technologies, enabling a broad understanding of concept visualization. VR allows you to see experience from different points of view and can also be used to generate 3D images of architectural related elements, such as a parapet detail, to assist in quick and deeper learning experience in understanding architecture and construction. Previously, the key argument inhibiting the presence of year in education of architectural design was the cost of purchasing technology related equipment and the difficulty of acquiring skills to operate it. Today, students are fluent in the application of CAD software and graphics processing applications. Students who have experience in computer modeling are becoming increasingly familiar with VR which helps promote changes in both education and architectural practice.

(Sardo et.al, 2018) [26] Augmented reality is discussed as a tool that immerses a person into an altered reality by not only connecting with the sight and sound, but also explores in incorporating the other three senses which have not been as explored as the other two. The multisensory AR is believed to involve a user to experience the better involved experience. The paper develops the interface by centering on a museum, but some of these can be added on to enhance AR experience. The interface system experiments also concentrate on moving from the big computer setups to more portable computer systems that can connect with the phones. Simulation for these senses is experimented with users to get results such as the simulation of touch and feel are explored with vibration, thermal simulation and air flow. There are not many options to stimulate the taste sense, but it has been studied that with the help of sight and smell, the sense for taste can also be developed. Vaporization was also a technique utilized to simulate smell. These experiments utilized some previously existing techniques and enhanced them to achieve the desired results. The conceptual model had a great response from the user survey even though it had some shortcomings and can still have better outputs with improvements.

(Megahed, 2014) [27] The use of AR is studied for the creative and perceptive development of the students in the field of design. The incorporation of these tools can shift the current teacher-oriented learning to a student-oriented learning. However, this incorporation still faces a lot of technological and social challenges. The pedagogy for the real life lessons has been in 2-D for a long time while the real objects itself exist in 3-D. Architectural design education is a field which requires a lot of practical knowledge and new methods to evolve the creativity of the students, with the constant changes coming up in the field of design it becomes essential to adapt, incorporating new techniques. The initial stage of data collection and analysis while experiencing the environment of different cities while sitting in the classroom gives a clearer perspective of context. The early stages of a design process involves most of the brainstorming, this stage requires the freedom for visualization and to be able to communicate it at its best. The next stage which incorporates a lot of reviewing and redo's can be made easy with the tools as mistakes can be corrected in real time on the spot. The paper discusses the working of two software in the field of education i.e., ARTHUR and AR Workbench.

(Abramovici, 2017) [28] Throughout the life-cycle of a building, maintaining the original conditions costs the most. The paper not only deals with visualizing the maintenance data but it also incorporates team wide communication and indicates any warning situations that require co-ordination. To make things efficient a co-ordination is developed amongst the machines and humans by using smart devices and internet of things. The whole concept is based on the graph model (GUI- Graph user interface) that analyses the type of goals and their approach to it. The model discussed is a cloud based system which requires the machines or objects requiring maintenance to be a part of the internet of things (cloud based system). The model connects with the individual technician leading to the activity that is related to other activities, the sensors record the results and values for follow up activities. The implementation of a prototype of the model was conducted to note the workability of the model, although the model could use more organized and a holistic planning. The author is now onto the stage of proof of concept.

4. DISCUSSION

AR and VR as tools for architecture and urban design have been theorized conceptualized since 1967, with the first research paper published by Schutz and North-Western University (1967)[2]. Since then, various advancements have been made and AR started to be used in indoor settings while VR was used for visualization. From then to now, the progress in the workability and application of these tools has been very slow paced, due to the reason that this technology has not been developed exclusively to suit this particular field. However in recent years, certain software's and hardware's have been developed for the specific use in the design field, which

explored immersive communication, real time visualization and data analysis. The scope of these tools has been explored more in concepts and experiments that work with the potential of the technology being used to (i) create an immersive user interaction, which creates a presence of your design and immerses you into the project to experience the impact of the project. (ii) To improve the real time positioning and tracking system. (iii) Provides the user to experience the design well before construction which enables to incorporate feedback. (iv) Incorporating these tools as part of an education system will expose the students to practical issues well before becoming proper architects. (v) Communicating designs to the workers is an issue that leads to a lot of causalities and management problems that can be life threatening, but can be inverted by using AR and VR to foresee these issues. (vi) Deteriorated/dilapidated heritage buildings with time could be experienced and visualized. (vii) Also, the complex urban fabric of the modern era, where all buildings are growing sky high and loosing individuality, becoming tedious to navigate around the city, for them these visual aided tool could work wonders.

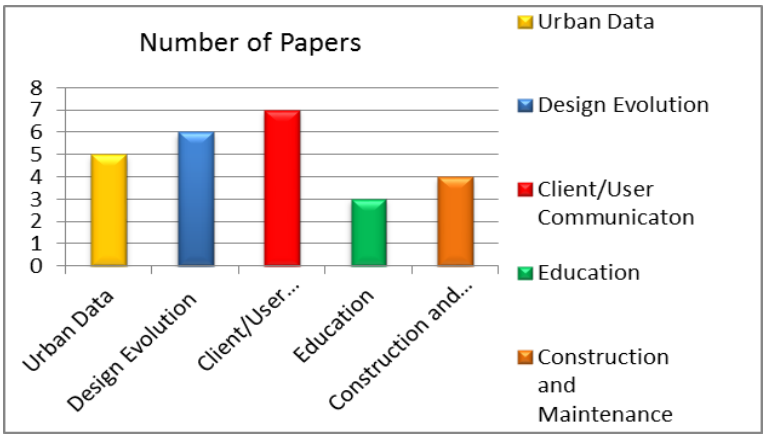

Figure 1. Categorization of AR/VR under diverse sub heads of architecture & urban design

However, these tools still come with a number of limitations in the field of technology and economics. The tools need to be more users friendly, accessible and economically viable to all the professionals and end-users. The industrial production of these tools can be seen taking place only after overcoming the limitations.

5. CONCLUSION

The article synthesizes the wide variety of architecture and urban design options that are being developed, conceptualized or is currently in use in the field of Augmented Reality (AR) and Virtual Reality (VR). The research papers briefly describe the diverse avenues that have the scope of development or improvement under AR. Collaboration and immersion are two aspects that are being explored in the field of design. With the complexities in design introduced by evolving time it becomes essential to manage this data in an efficient manner to seek maximum output. Reading, studying and sketching are important for understanding and designing for the people, but with AR and VR a fourth dimension is added to our comprehension. Human Computer interface is at the beginning of its era and it has the potential to change the depth of our understanding about the spaces and its ergonomics. Presently, the architects are already familiar with CAD tools and various simulation, analysis and management tools which typical VR technologies are allowing architectural design to be virtually represented and realised before construction. AR takes things to the next level and allows these virtual visualisations to be inserted into the physical reality. Thus, the ongoing future developments of AR and VR will play a significant role on improving the design of the built environment.

ACKNOWLEDGEMENT

I would like to place my thanks to Prof (Dr) Madhura Yadav, Director, School of Architecture and Design, Manipal University, Jaipur for her guidance. I also want to thank Dr. Vaneet Jishtu, Scientist, HFRI for help in motivation and guidance.

REFERENCES

[1] Husserl, E and Gibson, WRB 1931 Ideas: general introduction to pure phenomenology, G. Allen & Unwin Ltd., London. Pp. 465.
[2] Schutz, A and Northwestern University (Evanston Ill.) 1967 The phenomenology of the social world, Northwestern University Press, Evanston, III.
[3] Norman, D. 1988. The Psychology of Everyday Things, Basic Books, New York.
[4] Frazer, JH, Frazer, JM and Frazer, PA 1980 Intelligent Physical Three-Dimensional Modelling Systems, Computer Graphics 80 Conference, pp. 359–370.
[5] Seichter, Hartmut and Schnabel, Marc Aurel 2005 Digital and Tangible Sensation:
[6] Sørensen, Søren S 2006The development ofAugmented Reality as a tool inArchitectural and Urban design*Nordic Journal of Architectural Research*. Volume **19**, No 4, 2006.
[7] Döllner Jürgen, Konstantin Baumann, Henrik Buchholz 2006 Techniques and strategies of visualization complex 3D spatial data. *CORP 2006 & Geomultimedia06*. Assessed on 12.011.2020.
[8] Seichter, H 2007 Augmented reality aided design. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR. Retrieved from http://dx.doi.org/10.5353/th_b3896659
[9] Wang, Xiangyu (2009). Augmented Reality in Architecture and Design: Potentials and Challenges for Application. *International Journal of Architectural Computing*. 2009;7(2):309-326. doi:10.1260/147807709788921985.
[10] Abboud, Rana (2014). Architecture in an Age of Augmented Reality:Opportunities and Obstacles for Mobile AR in Design, Construction, and Post-Completion.
[11] Seichter, Hartmut 2020 Benchworks - Augmented Reality Urban Design. CAADRIA (2004). [*Proceedings of the 9th International Conference on Computer Aided Architectural Design Research in Asia* / ISBN 89-7141-648-3] Seoul Korea 28-30 April 2004, pp. 937-946.
[12] ARtoolkit. http://www.hitl.washington.edu/home/. Assessed on 15.11.2020.
[13] Kim, Mi jeong & Wang, Xiangyu & Zhu, Xingquan & Kang, Shih-Chung (2011). Augmented Reality Research for Architecture and Design. 10.4018/978-1-61350-180-1.ch013.
[14] Chung, DHJ, Zhiying, SZ, Karlekar, J, Schneider, M, & Lu, W (2009). Outdoor Mobile Augmented Reality for Past and Future On-Site Architectural Visualizations, Computer Aided Architectural Design Futures. pp.557-571.
[15] Wang, X, & Dunston, PS (2005). Real Time Polygonal Data Integration of Cad/Augmented Reality In Architectural Design Visualization, Computing in Civil Engineering (*American Society of Civil Engineers,ASCE*) (pp.1-8).
[16] Santos, P, Gierlinger, T, Stork, A & McIntyre, D (2007). Display and Renering Technologies for Virtual and Mixed Reality Design Review.*Interational Conference on Construction Applications of Virtual Reality*. pp.165-175.
[17] Belcher, D, & Johnson, B (2008). MxR: A Physical Model-Based Mixed Reality Interface for Design Collaboration, Simulation, Visualization and Form Generation.*Association for Computer Aided Design in Architecture*. pp. 464-471.
[18] SchubertGerhard, David Schattel, Marcus Tönnis, Gudrun Klinker, Frank Petzold (2015). Tangible mixed reality on-site Interactive augmented visualisations from architectural working models in urban design. São Paulo, Brazil : s.n., 2015. *The Next City - New Technologies and the Future of the Built Environment -16th International Conference*. pp 55-74 .
[19] Thabet, Walid & Shiratuddin, Mohd Fairuz & Bowman, Doug (2002). Virtual Reality in Construction. In: *Engineering Computational Technology*. pp. 25-52. Publisher: Saxe-Coburg Publications. March 2002, updated 2015.
[20] Milovanovic Julie, Guillaume Moreau, Daniel Siret, Francis Miguet (2017). Virtual and Augmented Reality in Architectural Design and Education: An Immersive Multimodal Platform to Support Architectural Pedagogy. *17th International Conference, CAAD Futures*, Gülen Çağdaş, Mine Özkar, Leman F. Gül and Ethem Gürer, Jul 2017, Istanbul, Turkey.

[21] Broschart, Daniel & Zeile, Peter (2014). Architecture – Augmented Reality Techniques and Use Cases in Architecture and Urban Planning. [https://www.researchgate.net/publication/269574525] Updated Feburary 2018.
[22] Redondo, Ernest & Navarro, Isidro & Riera, Albert & Fonseca, David (2011). Visual Interfaces and User Experience: Augmented Reality for Architectural Education: One Study Case and Work in Progress. *Communications in Computer and Information Science*. 166. pp. 355-367.
[23] Anke, Van Dessel (2019). The Use of Virtual Reality in Urban Planning Review paper on the (dis)advantages and a new perspective for future research. Master thesis - Master of Urbanism and Spatial Planning. Center for Urban Development, University of Antwerp.
[24] Ahmad K. Bashabsheh, Hussain H. Alzoubi, Mostafa Z. Ali (2019). The application of virtual reality technology in architectural pedagogy for building constructions. *Alexandria Engineering Journal*,Volume **58**, Issue 2. pp. 713-723. ISSN 1110-0168. https://doi.org/10.1016/j.aej.2019.06.002.
[25] Gębczyńska-Janowicz, Agnieszka (2020). Virtual reality technology in architectural education . *World Transactions on Engineering and Technology Education*. Vol. 18. No. 1. http://www.wiete.com.au/journals/WTE&TE/Pages/Vol.18,%20No.1%20(2020)/04-Gebczynska-Janowicz-A.pdf
[26] Sardo, J.D., Pereira, J.A., Veiga, R., Semião, J., Cardoso, P., & Rodrigues, J. (2018). A Portable Device for Five Sense Augmented Reality Experiences in Museums.
[27] Megahed. N. A. (2014). Augmented Reality Based – Learning Assistant for Architectural education.
[28] Michael Abramovici, Mario Wolf, Stefan Adwernat, Matthias Neges (2017). Context-aware Maintenance Support for Augmented Reality Assistance and Synchronous Multi-user Collaboration, Procedia CIRP, Volume 59; pp. 18-22

AUTHORS

Pearl Jishtu Presently pursuing her final year in Bachelor of Architecture, at Manipal University Jaipur. Finds motivation towards research work, in the role of technology in architecture. She has held positions of responsibility and has represented the University at National Association of Students of Architecture.

Prof (Dr) Madhura A Yadav An Architect-Planner with over 25-year experience. She is Founder Head of the School Architecture & Design and now Director of SA&D at Manipal University Jaipur. She is one of the founder members in establishing Faculty of Design at Manipal University Jaipur and founder head for fashion Design, Interior Design and Planning programmes. She is/was on several Committee/Expert Groups set up by Government of India., She is on Jury of various Architectural competitions. She is also on Editorial Board Architecture, Urban Design and Urban Planning India Journals. She has published number of papers in national and international journals and conferences. She is UNESCO certified Mentor for Community based participatory research. She is Indian Green Building Accredited Professional. She has conducted workshops/conferences on Waste Management, Bamboo Architecture, Earthquake Resistant Structures, Sustainable Cities/ buildings etc. Her current research interests are Architecture- Ecology interface and Eco- cities, Promotion of eco materials and Community Based Participatory Research. She is recipient of Distinguished Faculty in Architecture Award 2017 and Educational Leadership Award 2019.

AUTHOR INDEX