ISSN :25060

INTERNATIONAL JOURNAL OF SOFT COMPUTING & ARTIFICIAL INTELLIGENCE

VOLUME NO. 11
ISSUE NO. 1
JANUARY - APRIL 2023

ISSN :25060

International Journal of Soft Computing & Artificial Intelligence

Aim & Scope

IJSCAI provides a new forum for dissemination of knowledge on both theoretical and applied research on Artificial Intelligence and Soft Computing with an ultimate aim to bridge the gap between these two non-coherent disciplines of knowledge. The Journal offers survey and review articles from experts in the field, promoting insight and understanding of the state of art, and latest trends in the field. The content includes original research and innovation ideas, applications from all over the world.

All published papers are also available freely with online full-text content. This forum accelerates interaction between the above two bodies of knowledge, and fosters unified development in next generation computational models for machine intelligence. Special Issues devoted to topics of current interest will be published occasionally.

ISSN :25060

ISSN :25060

International Journal of Soft Computing & Artificial Intelligence

(Volume No. 11, Issue No. 1, January - April 2023)

Contents

ISSN :25060

PLANT DISEASE DETECTION USING DEEP LEARNING

[1]MUSAB PATEL, [2]ANAND SASANE,3 SHWETANK WALGUDE,4RUSHIKESH DHANE
Trinity College of Engineering and Research, Pune, India
E-mail: 3shwetank.walgude@gmail.com

ABSTRACT

Crop diseases are a major threat to food security, but their rapid identification remains difficult in many parts of the world due to the lack of the necessary infrastructure. The combination of increasing global smartphone penetration and recent advances in computer vision made possible by deep learning has paved the way for smartphone assisted disease diagnosis. Using a public dataset of 2000 images of diseased and healthy plant leaves collected under controlled conditions, we train a deep convolutional neural network to identify Cotton crop species and 3 diseases (or absence thereof). The trained model achieves an accuracy of 82.35% on a held-out test set, demonstrating the feasibility of this approach. Overall, the approach of training deep learning models on increasingly large and publicly available image datasets presents a clear path toward Web based crop disease diagnosis on a massive global scale. The advance and novelty of the developed model lie in its simplicity; healthy leaves and background images are in line with other classes, enabling the mode l to distinguish between diseased leaves and healthy ones or from the environment by using deep CNN . Novel way of training and the methodology used facilitate a quick and easy system implementation in practice. All essential steps required for implementing this disease recognition model are fully described throughout the paper, starting from gathering images in order to create a database, assessed by agricultural experts, a deep learning framework to perform the deep CNN training. This method paper is a new approach in detecting plant diseases using the deep convolutional neural network trained and fine-tuned to fit accurately to the database of a plant's leaves that was gathered independently for diverse plant diseases. The advance and novelty of the developed model lie in its simplicity; healthy leaves and background images are in line with other classes, enabling the mode l to distinguish between diseased leaves and healthy ones or from the environment by using deep CNN.

Key words- Demonstrating, Convolutional Neural Network, plant diseases

I. INTRODUCTION

Due to this Complexity and the large number of developed plants and their existing psychopathological conditions also fail, given trained agronomists and plant pathologists, to accurately diagnose specific diseases, leading to erroneous assumptions and solutions. Plant pests and rodents represent a major threat for the agriculture sector. In the world's agricultural sector, plant diseases cause significant production and economic losses. The In order to avoid more disruption, proactive action by detecting an early period of plant disease is one of the major challenges in overall crop disease control.Diagnosis of plant disease requires a significantly high degree of complexity through visual examination of the effects on plant leaves.

Research results suggest that climate change may modify stages and levels of pathogen production; it may also alter host resistance, which leads to physiological changes in host-pathogen interactions. The condition is further compounded by the fact that pathogens nowadays are quicker to spread worldwide than ever before. There can be emerging pathogens where they had previously been unidentified and where there is apparently no local knowledge to counter them. One of the basics of timely and proper treatment of plant diseases is of precision agriculture. It is Crucial to avoiding unnecessary expenditure

•of financial and other capital, while ensuring safer production, addressing the issue of long-term pathogen tolerance and minimizing the adverse effects of climate change. In plants, there are several ways to detect pathologies. Some diseases have no visible symptoms or the effect becomes apparent too late to act, and in those situations sophisticated analysis is mandatory.

Nonetheless, most diseases generate some sort of manifestation within the visible spectrum, so the Trained naked eye examination by a doctor is the main method that is used in use to detect plant disease. A plant pathologist should have good observational skills to accurately diagnose plant disease, so that signature symptoms can be identified.Signs of plant disease are evident in various parts of a plant; however, leaves are known to be the most commonly seen factor for detecting an infection. Hence, scientists have tried to automate the process of recognizing and classifying plant disease using vine images. Advances in artificial intelligence, machine learning, deep learning, image processing and graphical processing units (GPUs) can Expand and improve successful plant protection and plant development practices. Deep learning is about using artificial neural network architectures that include a fairly large number of computational layers. Convolution Neural Networks (CNNs) is the basic method of deep thinking used in the study. CNNs are one of the most effective tools of large-scale computer applications for modelling complex processes and pattern recognition, for example pattern recognition in images.

II. STATEMENT OF THE PROBLEM

The cotton plant is susceptible to several disorder (biotic and abiotic constraints) attacks due to temperature fluctuation, diseases, and pests. Indeed, the whole world produced nearly 576 kg per hectare of cotton crops, where only 10% of production loss occurred due to different cotton leaf diseases. The United States of America (USA) is a major exporter of cotton in the world and it obtained 5.1 billion US dollars in 2016, but there are well-known native pests which were the reason for the distraction of cotton farms. And, India has 24 percent of cotton land of the world and got 4.6 Billions of dollars in 2016, from which generally 18% of cotton crops' production was lost every year due to different diseases that attacked the cotton plants which had its impacts on losing almost nine hundred thousand of Indian rupees . Presently, in Ethiopia, nearly 12–15% of cotton crop plants are infected due to different diseases . In Ethiopia, performance evaluation of GTP-I showed that these diseases and pests are the main constraints of the world standards in cotton quality and quantity of production. This results in the downfall of the economy of both the farmer and the country detecting these diseases with bare eyes increased the complexity of cotton crops productivity which decreased the accuracy in identification precision. Even an expert would fail to assess and diagnose the diseases with their bare eyes, and this inadequate technique leads to more wastage of cotton crops. Due to these mistaken conclusions, most of the time, certain unnecessary pesticides which badly affect healthy cotton are applied. Leaving the farm for even a short time interval without production will affect overall nation GDP .

Deep learning incorporates image processing and data analysis as a path for more possible findings. As it has been a successful application, it has now entered the domain of agriculture. Today, several deep learning-based computer vision applications such as CNN (convolutional neural network), RNN (recurrent neural network), DBN (deep belief network), and DBM (deep Boltzmann Machine) are performing tasks with high accuracy. However, the most prominent application for this research work is CNN .

Nowadays, CNN techniques are used to detect different objects and to perform automatic drawings of instructions for analysis purposes. K-fold cross validation strategy recently recommended dataset splitting and boosted gen- eralization of the CNN model. Generally, the model de- veloped at the end was from scratch rather than any transferred learning model or pertained model.

Deep learning draws an attention in order to maximize the performances to classify different tasks which help to promise the human intervention data . In this real world, the usage of deep learning shows the major interest for decoding human brain activities. The problem is faced between intertrial and intersubject variability in electro encephalography signals, an indigenous access for attention-based bidirectional long-short-term memory. Convolutional neural network was analyzed among different factors that are classified into four classes of electro encephalography motor imaginary functions. Here, the usages of bidirectional long-short-term memory with the attention model accomplished the extraction of different features from the raw electro encephalography signals. Advancement of the clinical translation of the electroencephalography motor imaginary-based brain computer interface technology is applicable for varied request, where this system supports the paralyzed patients. The unusual achievements include the maximum accuracy and time resolved predictions. To make an efficient and effective interface system, the human plays an important role. Graph convolutional neural networks, a novel deep learning framework, addressed the issues in order to differentiate the four-class motor imaginary intentions by mutually agreeing through the similarity of electro encephalography electrodes. To find the motor imaginary, four tasks are preferred with the prediction of highest accuracy.

III. LITERATURE REVIEW

According to Shuyue , they outlined the different formats of graph convolutional neural network. It was prepared to process the uniform electro encephalography data for the purpose of predicting the four classes of motor imaginaries to relate with electro encephalography electrode. They addressed their data with the transformation of 2D to 3D perspectives. The structure was processed through these dimensional units.

A study stated that, in order to utilize the dynamic route of deep learning, they proposed short-term voltage stability. They managed the clustering algorithm to obtain short-term voltage stability to increase the reliability.In, it is stated that deep learning technique was applied to identify the leaf diseases in different mango trees. The researchers used five different leaf diseases from various specimens of mango leafs, where they addressed nearly 1200 datasets.

The CNN structure was trained with more than 600 images, where 80% are used for training and 20% are used for testing.

Remaining 600 images were used to find the accuracy and to identify the mango leaf diseases which showed the feasibility of its usage in real-time applications. The classification accuracy can be further increased if more images in the dataset are provided by tuning the parameters of the CNN model.

The research study states that the mechanism for the identification and classification of rice plant datasets are used to process the CNN model. For training, nearly 500 different images with diseases were collected for processing from the rice experimental field.

In, detection of cotton leafs were addressed with image processing. Here, K-means algorithms are used to segment the datasets.

The research showed the identification of diseases in banana plants which infect their leaf. In this research study, 3700 images were used for training, but there is no balanced dataset in each class. Researchers performed different experiments, for example, the training mode by using colored and grayscale image datasets and also by using different dataset splitting techniques. They obtained the best accuracy of 98.6% in colored image and 80% and 20% training to the validation dataset.

IV. SCOPE AND LIMITATION OF THE STUDY

This research study focused on developing an identification model for cotton leaf diseases and pests using deep learning technique called convolutional neural networking. Three common types of disease and pests such as bacterial blight,leaf miner, and spider mite have been affecting cotton productivity and quality. Also, the model applied made a supervised learning technique on datasets with four prime feature extraction process and 2400 datasets. The datasets are limited to four different feature descriptors. Taking into consideration the time constraints and reach of the regions that grow cotton, the research focused in the southern part of Ethiopia such as Arba Minch, Shele, and Woyto. MelkaWorer agricultural research center was also proposed as a focus area because it is responsible for cotton farms in SNNPR. Deep learning techniques were used to perform the automatic feature eradication from the different inputdatasets.

V. RESEARCH METHODOLOGY

This study used a design science to build and evaluate an approach that creates innovations and defines ideas, practices, technical capabilities, and products using qualitative or quantitative data. One of DSRM outputs is a model; it is a conceptual representation and abstraction of datasets. According to Hevner, Figure 1represents the processing model for this research.

Among different entry points, “problem-centered initiation” is the best fit for this design science research. The problem-centered initiation entry point is applicable because the problem is being observed by the researchers and business within the cotton disease identification domain. Figure 1depicts the DSRM proposed by the research study together with the activities adapted to this research.

VI. DATA COLLECTION AND SAMPLING TECHNIQUE

The sample leaf images which the researchers have used in this research are both primary as well as secondary types of dataset. Primary data is a type of data collected fresh for the first time. In this study, the primary types were collected from July to August 2019 from Arba Minch, Shele, and Woyto cotton farms where cotton plants are widely planted and there is high infection in SNNPR, whereas secondary data collected in each class were obtained from Melaka- Worker agricultural research center founded in the Afar region and SNNPR.

For this study, the researcher has used purposive or judgmental sampling techniques, selecting three infected and a healthy sample from the population, which is nonprobabilistic.

During data collection, 2400 images of data are captured and distributed into four equal classes such as bacterial blight, healthy, leaf miner, and spider mite used to train with balanced dataset, as shown in Figure 2.

6.1. Cotton Images' Sample Digitization. The data acquisition system in this research was used with regard to generate clear, unbiased, and simplified digital images of leaf in the cotton plant sample database for further analysis and processing.

The aim was to provide the digitizing system with uniform lightning or balanced illumination. The images captured using a smartphone camera and digital camera are then transferred to a computer, displayed on a screen, and stored on the hard disk in the PNG format as digital color images.

6.2. Image Data Preprocessing Inserting preprocessed images into a network is the first and basic task in all image processing projects. Common image preprocessing tasks in any image processing project are

vectorization, normaliza- tion, image resizing, and image augmentation.
In this re- search, these image preprocessing tasks are carried out before going to further deep learning processing using OpenCV library in python .Data augmentation is also used to generate more training datasets from the real sets for data samplings.

6.3. Feature Extraction. Deep learning solves different shortcomes of machine learning feature extraction such as extracting features manually by using the best and robust technique called a CNN . The layers are used to learn the knowledge. With the use of filtering mechanism the data are used to match and extract their values.

6.4. Dataset Partitioning and Model Selection Methodology. The used dataset partitioning technique is K fold cross- validation which is partitioned as K values, where K + 1 have to be obtained for the upcoming divisions. For this research, the study researcher has assigned the K value as 10 because it is recommended for deep learning [8, 20].

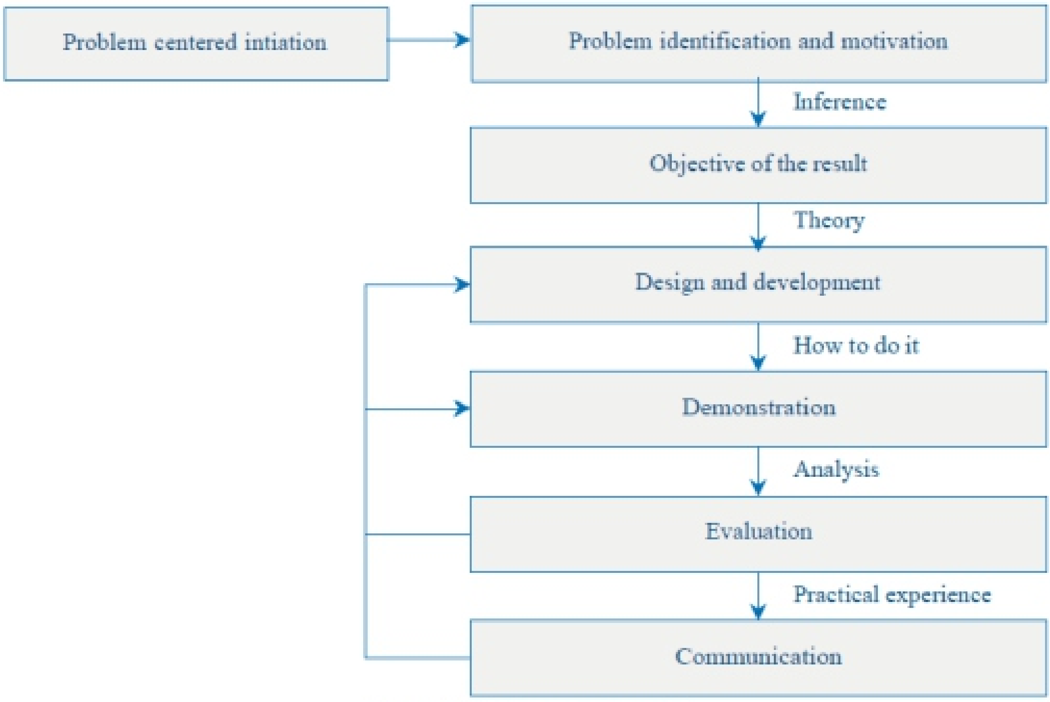


Figure 1: DSRM processes' model.

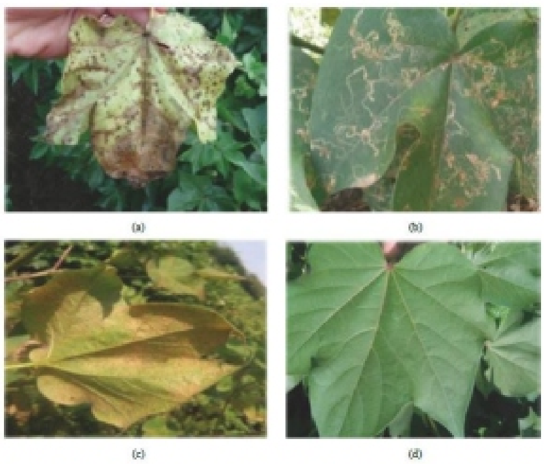


Figure 2: Dataset classes: (a) bacterial blight, (b) leaf miner, (c) spider mite, and (d) healthy.

This routine activity, 80% (2160 leaf images) yield the most appropriate performance which are trained and rest 20% (240 leaf image) are used for testing; thus, the system was validated.

*6.5. Tool Selection.*To collect cotton leaf images for this research, two image capturing devices were used such as a smartphone and digital camera. The proposed model was implemented using python version 3.7.3 for its usages. Also, the model is trained on the deep learning package called Keras, Version: 2.2.4-tf TensorFlow backed. TensorFlow, Version: 1.14.0 was recommended to adopt the proposed system. To evaluate the performance, many experimental setups were conducted with the help of a graphical user interface using Tkinter. From hardware, training and test was carried out on CPU instead of GPU.

*6.6. Evaluation Techniques.*To evaluate the routine of the structure, the researchers used various techniques in different periods, such as the developmental stage and at the end. First, the researchers evaluate the acquirements of the prototype using the confusion matrix and four evaluation metrics for confusion matrix reports such as F1-score, Precision, Recall, and Accuracy on the test dataset. Secondly, in this study for subjective evaluation, the researcher has used a questionnaire to measure the performance of a prototype by domain experts, as shown in Figures 3and 4. An objective evaluation has been made using the experimental analysis to test an artifact. Finally, the result of the evaluation depicts the practical applicability of the model.

VII. DESIGNING OF COTTON PLANT DISEASE AND PEST IDENTIFICATION MODEL

The first task in this model designing is image acquisition from the field with digital camera and smartphone. Then, image preprocessing techniques were applied to prepare acquired images for further analysis. After this, preprocessed images were inserted into the CNN algorithm to feature extraction with neural network. Then, best-suited extractions to represent the image are extracted from the image using an image analysis technique. Based on the extracted features, the training and testing data that are used to identify are extracted. Finally, a rained knowledge base classifies a new image into its class of syndromes, as shown in Figure 5.

VIII. THE ARCHITECTURE OF CNN FOR THE MODEL

CNN architecture consists of two broad sections such as feature learning and classification section. In general, the cotton images feed into an input layer and end with an output layer. The hidden layer consists of different layers, as shown in Figure 6. Here, a cotton leaf and the output will be the class name of such an image also called the label of cotton leaf diseases or pests. In general, for this proposed architecture, each cotton leaf images with addition of neurons are augmented with considerable weights.

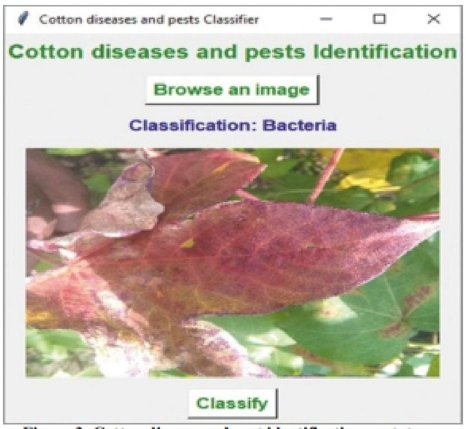


Figure 3: Cotton disease and pest identification prototype identified as bacteria.

Figure 4: Cotton disease and pest identification prototype identified as spider mite.

Output of the augmentation process to the upcoming layers are processed and duplicated to next layer. Output layers show the prediction tasks for calculating neurons for this research.

IX. EXPERIMENTAL RESULTS

During experimentation, different experiments were undergone to get an efficient model by customizing various parameters that provided different results. Those parameters are dataset color, number of epochs, augmentation, optimizer, and dropout. According to Serawork Wallelign [19], augmented RGB colored images provided about 15% improvement on accurate than that of not augmented.For this new model, the researcher has trained three different numbers of epochs such as 50, 100, and 150.However, the model achieved the best performance on 100 epochs, as shown in Figure 7. Nitish Srivastava [5] added a dropout in the CNN given additional performance (2.7%).

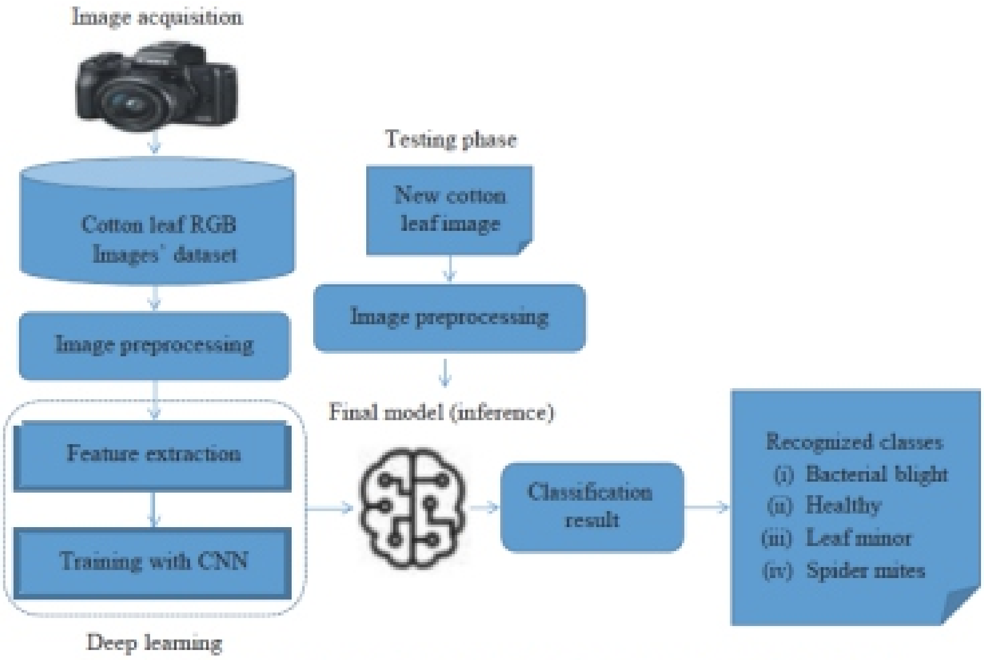


Figure 5: Cotton leaf diseases and pests recognition model process

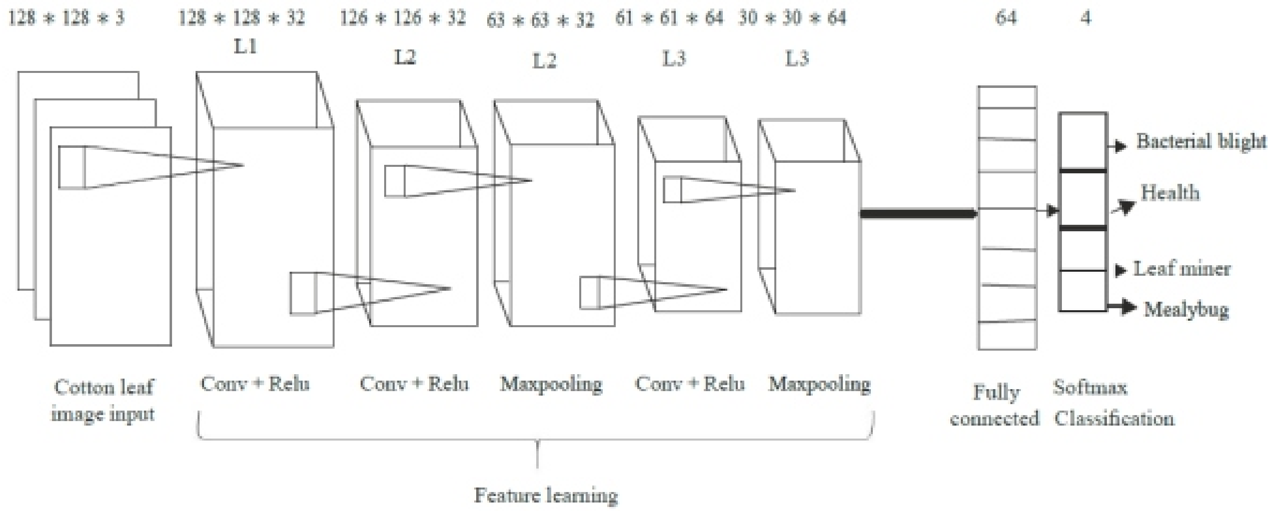


Figure 6: Developed CNN architecture for training

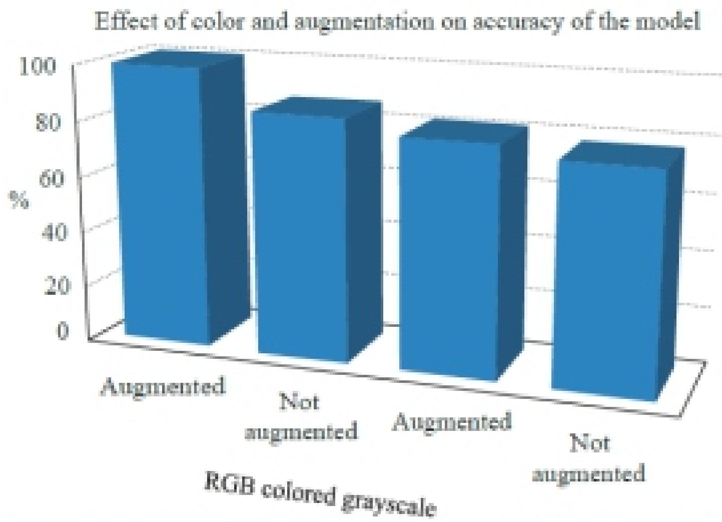


Figure 7: Color and augmentation parameters' experiment result.

Therefore, during the experiment, the researcher used 0.25 and 0.5 dropout percent in each layer and achieved the best performance in 0.5 dropout percent. Finally, a very important experiment was carried out on the regularization method that optimization algorithms' usage could minimize the loss through iterations by updating means according to a gradient. From Figure 8, it is observed that the effects of numbers on epochs and regularization methods are identified. For this research, two most recent and used optimization algorithms are used such as RMSProp and Adam, but the Adam optimization algorithm reduces loss by 2.5%, as shown in Figure 9.

Researchers observed highest training accuracy at the 100th epoch as 0.990. The graphs show all the training and validation success rates that the network achieved during the process, as shown in Figure 9, and the loss graph is shown in Figure 10.

X. RESULTS AND DISCUSSION

To analyze the performance of the model, the last result is achieved using parameters such as K-fold cross-validation using 10 folds. RGB-colored image dataset with augmentation provides 15% best performance for the model. The researchers used the transferred learning CNN model and the grayscale dataset achieved 98.6% accuracy [6]. However, color is the main and most decisive feature in cotton detection and classification; therefore, using a colored dataset takes a long time to train the model to add performance even if it is a complex layer. The number of epoch with 100 iterations and the Adam optimization method is very significant to boost the model performance by 10% and 5.2%, respectively. In the end, this developed CNN model achieves 98% of bacterial blight, 94% healthy, 97.6% of leaf minor,and 100% of spider mite, which are correctly classified.

Additionally, the researcher has used different preprocessing techniques for noise removal. The main factors for the misclassification of the result exist between bacterial blight, healthy, and leaf miner. The overall performance of the model, as shown in the confusion matrix, is 96.4% accurate for diagnosis of leaf disease and pests of cotton plants.

XI. PROTOTYPE DEVELOPMENT AND EVALUATION

For the prototype, the researchers focused on the convention of the digital forensic investigation process, which is ISO and IEC to evaluate the prototype in terms of efficiency, effectiveness, fault tolerance, helpfulness, learn ability, and the control to assess the quality of the prototype. For the time being, the system prototypical test is carried out as a desktop application which is conducted with the help of Tkinter, a graphical user interface in Python programming language.

For questioners, evaluators were allowed to rate the options as extremely satisfied, very satisfied, somewhat satisfied, not so satisfied, and not at all satisfied for five closed-ended questions and one open-ended question.

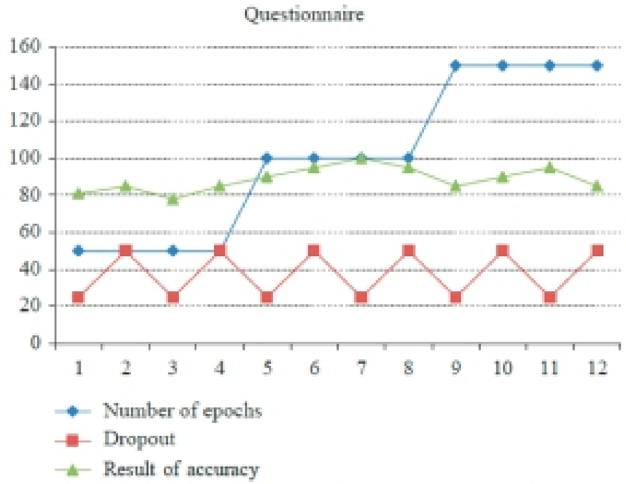


Figure 8: Effects of number of epochs and regularization methods during experiment.

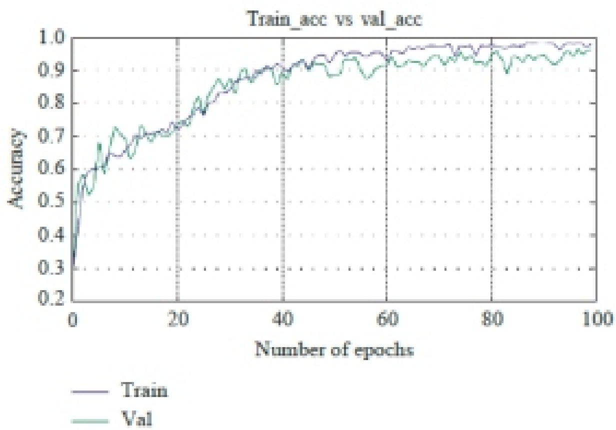


Figure 9: Training accuracy and validation accuracy of the model

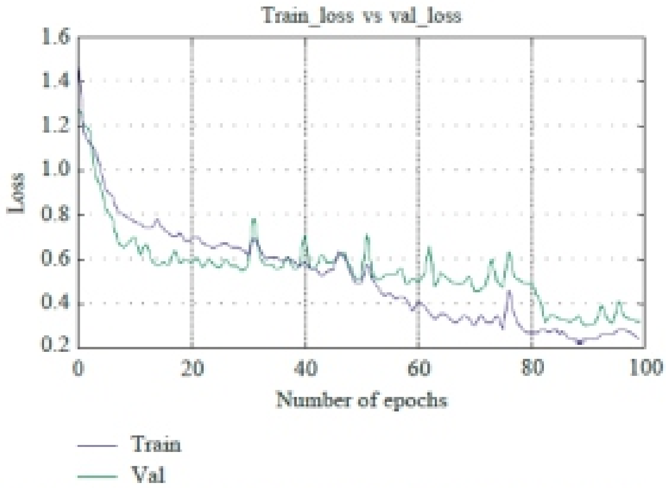


Figure 10: Training loss and validation loss of the model

The questionnaires are distributed to Ethiopian cotton farm experts, as shown in Figure 11. The data obtained from the farmers are recorded in Table 1

XII. CONCLUSION

This deep learning-based model was implemented using Python and Keras package, and Jupyter was used as a development environment. Different experiments have been undergone in this research study to get an efficient model by customizing various parameters such as dataset color, number of epochs, augmentation, and regularization methods. RGB-colored image dataset with augmentation provided 15% best performance for the model. The numbers of epoch and regularization methods are very significant to boost the model performance by 10% and 5.2%, respectively.

The proposed prototype has achieved the highest efficiency of 96.4% for identifying each class of leaf disease and pests in cotton plants. Developments of such automated systems are used to assist the farmers and experts to identify cotton disease and pests by leaf visual symptoms. Obtained results evidence that the designed system for the farmers are much helpful in order to reduce the complexity, time, and cost of diagnosing the leaves from any diseases.

FUTURE WORKS

The main challenge while developing an object detection model on deep learning was to collect a large number of training high-quality images with different shapes, sizes, different backgrounds, light intensity, and orientations in different classes. Therefore, future researchers should try to include a solution for such challenges in their work and not only identify but also suggest remedies for diseases and pests.

Ethiopia launched the satellite in 2019, and this is the best initiative for the future researcher to access remote-accessing high-resolution satellite images to train high-performance deep learning technique-based model.

DATA AVAILABILITY

During data collection, 3117 images of data are collected from those varied environments.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

ACKNOWLEDGMENTS

This research was funded by the Arba Minch University, affiliated with Ministry of Science and Higher Education, Ethiopia.

REFERENCE

[1] V. K. Vishnoi, K. Kumar and B. Kumar "Plant disease detection using computational intelligence and image processing", Journal of Plant Diseases and Protection Springer, pp. 19-53, 2021.

[2] Mingjing Pei; Min Kong; Maosheng Fu; Xiancun Zhou; Zusong Li; Jieru Xu Application research of plant leaf pests and diseases base on unsupervised learning

[3] Xulang Guan Brownell Talbot School, Omaha NE, US A Novel Method of Plant Leaf Disease Detection Based on Deep Learning and Convolutional Neural Network

[4] Plant disease classification using deep learning Akshai Kp Karunya Institute of Technology and Sciences, Coimbatore,India

[5] CSE-APEX Chandigarh University Gharuan Mohali, Punjab, India

*[6] Review on Plant Disease Detection using Deep Learning:**Banushruti Haveri** Affiliation Department of Electronics and Communication, Dayananda Sagar College of Engineering, Bangalore **K. Shashi Raj** Affiliation Department of Electronics and Communication, Dayananda Sagar College of Engineering, Bangalore*

*[7] Comparison of Pre-Trained Models Using Transfer Learning for Detecting Plant Disease : **Bincy Chellapandi** Affiliation Department of Information Technology VESIT Mumbai, India **M. Vijayalakshmi** Affiliation Department of Information Technology Vivekanand Education Society's Institute of Technology Mumbai, India **Shalu Chopra** Affiliation Information Technology VESIT, Mumbai, India*

[8] The Plant Disease Detection Using CNN and Deep Learning Techniques Merged with the Concepts of Machine Learning Sajitha N;Sourabh Nema;Bhavya K R;Parthasarathy Seethapathy;Kumud Pant

[9] A Performance-Optimized Deep Learning - Based Plant Disease Detection Approach for Horticultural Crops of New Zealand Muhammad Hammad Saleem;Johan Potgieter;Khalid Mahmood Arif

[10] A Hybrid Model for the Classification of Sunflower Diseases Using Deep Learning Akash Sirohi;Arun Malik

[11] Rose Plant Leaves: Disease Detection and Pesticide Management using CNN Authors Mohammad Ibrahim Khaleel;Paineni Gowtham Sai;Adabala Uday Raghavendra Kumar;Purushothamman Raja;Vinh Truong Hoang

[12] Progress on Deep Learning Models for Plant Disease Detection: A Survey Adedamola O. Adedoja;Pius A. Owolawi;Temitope Mapayi;Chunling Tu

[13] Review on Leaf diseases detection using Deep learning Shweta Bondre;Ashish K. Sharma

[14] Plant Disease Detection Using Advanced Deep Learning Al gorithms: A Case Study of Papaya Ring Spot Disease Prajwal Maski;Asokan Thondiyath

[15] Plant Disease Detection based on Deep Learning Approach Faizan Akhtar;N. Partheeban;A. Daniel;Srinivasan Sriramulu;Saloni Mehra;Nishant Gupta

[16] Convolutional Neural Networks based Classification and Detection of Plant Disease P. Vishnu Raja;K. Sangeetha;Ninisa B A;Samiksha M;Sanjutha S S

[17] Leaf Disease Detection of Multiple Plants Using Deep Learning Shital Pawar;Sakshi Shedge;Nibedita Panigrahi;A P Jyoti;Pradnya Thorave;Samina Sayyad

[18] Plant Leaf Disease Detection using Deep Learning: A Review Rina Bora;Deepa Parasar;Shrikant Charhate

[19] Tomato And Potato Leaf Disease Prediction With Health Benefits Using Deep Learning Techniques K. Karthik;S. Rajaprakash;S. Nazeeb Ahmed;Rishan Perincheeri;C. Risho Alexander
[20] Detection of Cotton Plant Disease for Fast Monitoring Using Enhanced Deep Learning Technique Himani Bhatheja;N. Jayanthi

ISSN :25060

FACE MASK DETECTION USING MACHINE LEARNING

[1]SHRUTI MISHRA, [2] VAISHALI JADHAV
[1]Student, St. Francis Institute of Technology/Mumbai University
[2]Asst. Professor, St. Francis Institute of Technology/Mumbai University
E-mail: 1shrutimishra@student.sfit.ac.in, 2vaishalijadhav@sfit.ac.in

A B S T R A C T

Coronavirus disease 2019 has affected the world seriously. Coronavirus belongs to the class of viruses A coronavirus identified in 2019, SARS-CoV-2, has caused a pandemic of respiratory illness, called COVID-19. COVID-19 is the disease caused by SARS-CoV-2, the coronavirus that emerged in December 2019. COVID-19 can be severe, and has caused millions of deaths around the world as well as lasting health problems in some who have survived the illness. Coronavirus spreads through droplets and virus particles released into the air when an infected person breathes, talks, laughs, sings, coughs or sneezes. After sneezing or coughing the large droplets that fall into the ground causes the viruses to accumulate at one places and this leads to transfer of disease from person to person. Wearance of mask prevents the spread of such infections and disease. Mask causes the hinderance to the virus to enter through the nose and throats through that accumulated place that resulted into the deposition of virus from sneezing or coughing. Due to this the govt has made mandatory to follow the proper wearance of the mask in public. But achieving this practice manually, it is a very tedious task. Therefore, this demands the existence of automated face mask detection system, that identifies automatically whether the person has wore a mask or not. The detector system should be viable and has to deployable in public so as to curb the spread of disease thereby making the public to mandatorily wear the mask. In this paper we aim to perform a comparative analysis of various sophisticated image classifiers based on deep learning, in terms of vital metrics of performance to identify the effective deep learning based model for face mask detection.

Key Words—Facemask Detection, RMFD, SMFD VGG-16, MobileNetV2, InceptionV3.

I. INTRODUCTION

Coronaviruses belongs to the class of viruses that can cause respiratory diseases in humans. Severe acute respiratory syndrome (SARS), Middle East respiratory syndrome (MERS) and the common cold are examples of coronaviruses that cause illness in humans. The new strain of coronavirus — COVID-19 — was first reported in Wuhan, China in December 2019.

The virus has since spread to all over the world. Millions of people has been infected As of this writing (10/11/2021), more than 196,910,000 people in the world have been infected. Over 4,850,000 people have died.COVID (2020) Some 192 countries and territories on all continents have now reported cases of COVID-19. COVID-19 enters your body through your mouth, nose or eyes (directly from the air-borne droplets or from transfer of the virus from your handsto your face). The virus travels in respiratory droplets released into the air when an infected person coughs, sneezes, talks, sings or breathes near you (within 6 feet). One may be infected if you inhale these droplets. This makes very important for the person to wear a mask on the regular basis in order to avoid the direct contact from the person droplets and also from the infected patients. Face mask detection is detecting whether a person is wearing a mask or not. Face detection detects the presence of face in real-time videos or in any photos. Several rules are set to force wearing a facemask in public and work places, which represent hotspots for the spread of this infection. However, not every individual is aware or compliant, thus risking his or her life and the lives of others by not wearing a mask. Real- time monitoring of facemask wearing for a large group of people is becoming a difficult task. Manual monitoring is in general hard to enforce because of the manpower needed to efficiently protect public spaces and to ensure that individuals are wearing masks correctly. Aside from the cost problems and managerial effort, the biggest problem is the health factor because a certain set of employees will be in contact with hundreds of people daily, which poses a risk of them acting as infection points, therefore we aim to eliminate the human factor contact.[7] Recently, deep learning (DL) has been used in many do- mains and solved many complex problems, providing, therefore, significant results. DL allows analyzing and interpreting massive volumes of data in a fast and accurate way. Therefore, in this paper, we are proposing an approach that would help enforce the facemask policy and monitor it with ease in real-time videos. Theproposed system will help the authorities and commercial spaces monitor facemasks easily and efficiently. The novelty of this paper with regard to exiting works is proposing an efficient and accurate approach for real-time videos. The proposed approach provides accurate detection of facemask wearing and whether it is worn in an appropriate way or not in real-time. To do this, a complete dataset is collected using public datasets and also our own one.For face mask detec- tion we have used Haarcascade classifier that bifurcates the image into positive (contains mask) or negative(do not con- tain mask). To be able to do this accurately, the algorithms are trained on huge datasets (RMFD- Real Masked Face Dataset) containing 5000 mask faces and 90,000 normal faces of 525 people after that also used (SMFD- Simulated masked face recognition datasets).[19] This proposed system can be integrated with systems installed on the various places, that allows the person to have access on that place. This approach has the advantage to be fast and suited to edge devices, and it provides excellent results for object detection. Performance metrices that has been used are, Accuracy, Precision, Re- call, F1 score, confusion matrix, and classification report. The proposed solution can be implemented in real-world surveillance cameras in public areas to check if people are following rules and wearing marks. The solution can be easily implemented with minimum resources.

II. RELATED WORKS

W., Carin L., Dzau V. Ting (2020) propose a portable face mask service stage detection system. It is based on the texture features of face masks and the stage of the mask to which it belongs to. The main contributions of this paper are threefold. It finds out the correct stage of the mask through textual features of the mask. The main contributions of this paper are threefold. Accuracy : 82%. There are many research papers that have proposed a solution for the detection of face mask on the human face to curb the spread of corona virus. In this paper we have mainly focused on perform- ing the comparative study on the pre-trained model and to conclude the best performing pre trained model for the face mask detection.GagandeepKaur ,RiteshSinhaKaur, Sinha, Tiwari, Yadav, Pandey, Raj, Vashisth and Rakhra (2021) have performed the implementation on face mask detection using CNN. They have provided a simple way to achieve this objective utilising some fundamental Machine Learning tools as TensorFlow, Keras, OpenCV and Scikit-Learn. The suggested technique recognises the face in the image or video and then determines whether or not it has a mask on it. The implementation provides an accuracy of 83%. Shashank D. JoshiSandesara, Joshi and Joshi (2020) has proposed a face detector model that uses stacked Conv2D model. They have used this convolutional neural network to deduce the minutes of pixel in an image. The proposed model is a stack of 2-D convolutional layers with relu activations as well as Max Pooling and they have implemented this model by using Gradient Descent for training and binarycross-entropy as a loss function. They have used a RMFD dataset from kaggle. Validation or test accuracy is 95% has been achieved in this model. The time taken for the training of model is quite high in this case. In order to reduce the training time Mohamed LoeyGunasekaranManogaran- Mohamed Hamed N. TahaNourEldeen M. KhalifaLoey, Manogaran, Taha and Khalifa (2021) proposed a hybrid model that consisted of 2 components; The first component is designed for feature extraction using resnet, the inculcance of resnet has caused the fast training of model since it uses the transfer learning approach, while the second component is used for the classification process of face masks using decision trees, Support Vector Machine (SVM). 3 dataset were used in this case RMFD, SMFD, LFW. GayatriDeorel, Ramakrishna Bodhula, Dr. VishwasUdpikar, Prof. VidyaMoreMangayarkarasi, Anusha and Pranavikha (2021), has implemented a automated video survillnce system in such a way that uses certain physical parameters of the face and performs the detection of mask automatically through video installed in a public place. In order to make it more accurate they estimated the distance from camera, eye-line detection, facial part detection. This unique approach for the problem has created a method simpler in complexity thereby making real time implementation feasible. The model has achieved a accuracy of 89%. It is very important to keep check on the quality of mask the person is using in order to effectively curb the spread of virus, Yuzhen Chen, MenghanHu , Member, IEEE, Chunjun Hua, GuangtaoZhaiChen, Hu, Hua, Zhai, Zhang, Li and Yang (2021)aimed at solving the problem that we do not know which service stage of the mask belongs to, we propose a detection system based on the mobile phone. They first extract four features from the gray level co-occurrence matrixes (GLCMs) of the face mask's micro- photos. Next, a three-result detection system is accomplished by using K Nearest Neighbor (KNN) algorithm. The result achieved was 82%.

III. PROPOSED SYSTEM

The aim of this paper is to compare different transfer learning models and to evaluate their performance for the detection of presence of face mask. A Convolutional Neural Network (CNN) convolves the input images or feature maps with convolution kernels in order to extract higher-level features. Thus it is an effective tool for Computer Vision tasks like image classification, object detection, patterns identification, etc.[1] The neural network architectures studied, evaluated and compared in this paper are VGG16, and MobileNetV2. Classification of the image is better achieved by the transfer learning of these models.

In the first comparison proposed model uses VGG-16 for better classification. Fully connected layer which is the top most layer of the existing model has been removed and replaced with flatten layer, dense layer and dense softmax layer for the improvement in classification Overfitting is prevented by using the drop out for dropping out certain values at the random. Softmax layer is used for multipleclassification of the facial features in an image. Except for the last layer, the ReLu activation mechanism has been used in all layers. The number of images used in the training phase 75% of the total number of images in the dataset, with the remainder used for confirmation.[8]

In second comparison, the proposed model use Mo bileNetV2 a Deep Neural Network, has been deployed for the classification problem. Weights that has been pre-trained taken from imagenet has been loaded through tenserflow. Then the base layers has been freezed to avoid disfigurement of already learned features. Followed by addition of new trainable layers, these layers are trained on the dataset that has been used in the implementation so that it can determine the features to classify a face wearing a mask from a face not wearing a mask. Then finetuning is performed on the model, and then the weights are saved.[5]

In third comparison, proposed model uses InceptionV3 Neural Network, has been deployed for the classification problem. Pre-trained weights of InceptionV3 is used that is already trained on Imagenet dataset and the last layer is being removed i.e softmax layer and extraction of 2048 feature vector is done. After extraction of the feature vector 2 more dense layers are added of 128 neurons to build fully connected architecture to classify the image into masked and non-masked images and lastly trained on mask dataset. [25]

IV. METHODOLOGY

In this section, a classifier is being made that can differentiate between faces with masks and without masks. So for creating this classifier, we need data in the form of Images. We have a dataset containing images faces with mask and without a mask. We have fine-tuned a pre-trained net- work called MobileNetV2. [6] ,VGG-16 and ImageNetV3 which is trained on the Imagenet dataset with our neural network model. Initially we have read all the images and assigned them to some list. All paths are received associated with that of image and that has been labelled accordingly. Our dataset contained two folders viz- with_masks and without_masks. So we can easily get the labels by extracting the folder name from the path. Also, we pre-process the image and resize it to 224x 224 dimensions. The next step is to load the pre- trained model and customize it according to our problem statement. So only the last layer has been manipulated i.e. output layers and few layers have been added by own. [24] Now we need to convert the labels into one-hot encoding. After that, we split the data into training and testing sets to evaluate them. Also, the next step is data augmentation which significantly increases the diversity of data available for training models, without actually collecting new data. Data augmentation techniques such as cropping, rotation, shearing and horizontal flipping are commonly used to train large neural networks. The next step is to compile the model and train it on the augmented data. Next we create CNN model. This convolution networkconsists of three pairs of Conv and MaxPool layers to extract features from the dataset. Which is then followed by a Flatten and Dropout layer to convert the data in 1D and ensure over fitting. [30] Now that our model.

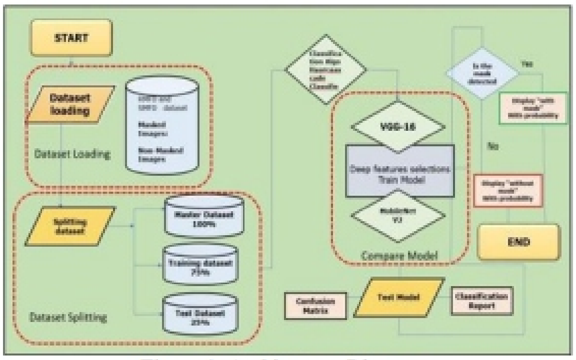


Figure 1: Architecture Diagram

is trained, we can modify the code in the first section so that it can detect faces and also tell us if the person is wearing a mask or not. In order for our mask detector model to work, it needs images of faces. For this, we will detect the frames with faces using the methods as shown in the first section and then pass them to our model after pre- processing them.[5] The next step is to find the faces in our frames. The faces variable contains the top-left corner coordinates, height and width of the rectangle encompassing the faces, we can use that to get a frame of the face and then pre-process that frame so that it can be fed into the model for prediction. The pre-processing steps are same that are followed when training the model in the second section. After getting the predictions, we draw a rectangle over the face and put a label according to the predictions.[23]

V. PROCESS OF A SYSTEM

This section describes the process flow of a system from data collection, converting the raw data to pre processed format, face detection using Haar Cascade classifier and it also discusses the 3 pre-trained model that have been used for training the images of the dataset and lastly the metrics based on which the performance of the system is being evaluated.

5.1 Data Collection

We have used a two different dataset for face mask detection that in total consists of 1339 photographs. Real time images of 120 photographs were included. For detecting masks from video used CCTV footage and Webcam, both of the photos are in RGB. To avoid overfitting, we collected data from different datasets in which we have taken the reference of Real-World Masked Face Dataset (RMFD) [28] and the Simulated Masked Face Dataset (SMFD) [29], which we used for training and testing purpose. (RMFD) and (SMFD) dataset.

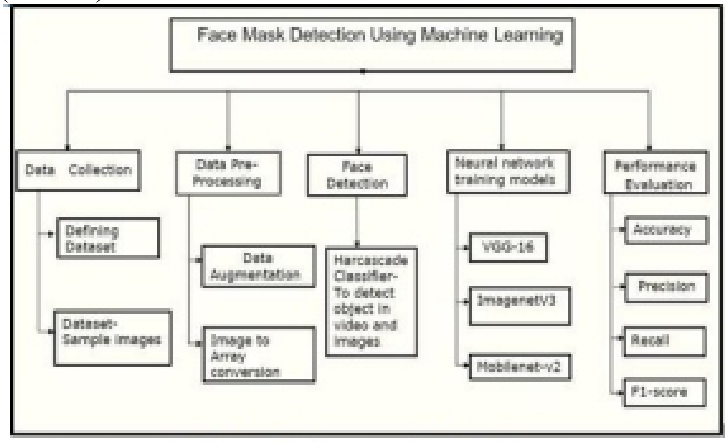


Figure 2: Process of a System

It contain multiple masks and non-masks of the same person Face images. For this purpose, we have established two face recognition sample sets for masks.

(1) **Real mask face recognition data set:** After sorting, cleaning and labeling a sample from the Internet, it contains 5,000 mask faces and 90,000 normal faces of 525 people. [14]

(2) **Simulated masked face recognition datasets:** We put on the masks on the faces in the public face datasets, and obtained the simulated masked face dataset of 500,000 faces of 10,000 subjects.The face identification accuracy on the dataset is over 94%. [24]

Figure 3: Sample Image of the Dataset

5.2 Data Pre-Processing

In machine learning, in order to make it understand to the machine and to make it machine readable, data pre- processing is performed to convert the raw data into processed format. In order to make the process easier we have divided the data-pre-processing into different number of stages. Here we have performed Data augmentation and conversion of image to array as a part of data pre-processing stage. Data Augmentation-Technique to increase the diversity of your training set by applying random (but realistic) transformations, such as cropping, rotation, shearing and horizontal flipping are commonly used to train large neural networks.Image to Array Conversion –Images to array conversion have been performed in order to train the deep learning model based on the features of the input images from the webcam. The library used is NumPy for the conversion of image to array. [20]

We have resized it to 224x 224 dimensions. For Image to Array Conversion Numpy is being used.

5.3 Face-Detection

Face detection – also called facial detection – is an artificial intelligence (AI) based computer technology used to find and identify human faces in digital images. For face detection we used the OpenCV based Haar cascade classifier which was developed by the Paul Viola and Michael Jones Nieto-Rodriguez et al. (2015). Haarcaascade classifier is an object detection algorithm that is used to identify faces in real time system. It is provided with lot of positive images with faces and lot many of negative images that doesn't consist of any faces to train. Initially a haarcascade face.xml file is loaded followed by loading of greyscale pattern of input image. This causes to happen a face throughout the image. If features are identified , the co-ordinates of that region is calculated as Rect(x, y, w, h). If we have those locations we are able to retrieve the ROI of the image. Lastly on the calculated ROI, a classifier been applied to detect if the person is wearing a mask or not.[15]

5.4 Neural Networks

CNN - One of the main parts of Neural Networks is Convolutional neural networks (CNN). CNN is made of several layers in order to perform the training and classification. Those layers are fully connected layer that means each layer has neurons that is connected to the neurons of the other layer. It performs certain operations such as pooling and Relu in order to decrease the weights and to prevent the exponential growth in the computation required to operate the neural network.

Limitations of CNNA Convolutional neural network is slower in operation due to existence of pooling operations If the CNN has several layers then the training process takes a lot of time if the computer doesn't consist of a good GPU. A ConvNet requires a large Dataset to process and train the neural network.[30]

5.4.1 Transfer Learning

Supervised machine learning model comprises of labelled data through which the model is trained and through which the machine predicts the output. Labelled data means some input data is already tagged with the output. Training data provided to the machines earlier acts a supervisor that helps them to predict or classify the results accordingly. In supervised machine learning it takes a lot of time to train the model due to large number of back propagations. Epochsare very high followed by time taking factor.Gazzah and Bencharef (2020)

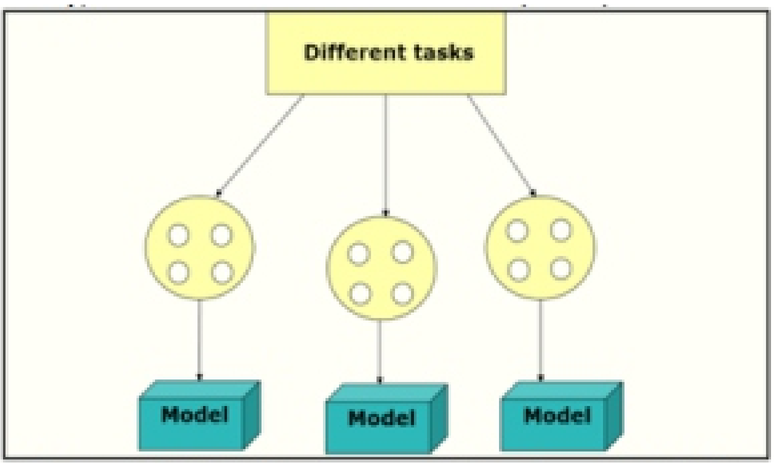


Figure 4: The traditional supervised learning setup in ML

If we want to train a model that classify a person as with mask or without mask, we require to re-built the model from scratch. Since training of a model from scratch requires a lot of time and money, many problems are not profitable with the traditional machine learning approach. It will also require a vast amount of dataset of masked and unmasked images to generate a meaningful results. These data requirements are often difficult to satisfy in practice.Transfer learning involves the usage of pre-trained model that is already been trained on 1000 no classes and in millions of images and also on vast varieties of dataset. These features do not require the model to be trained from scratch. It also requires comparatively lesser no of time then traditional machine learning approach. It also has less number of iterations or epochs. It uses the previously stored knowledge while training. Inspite of learning new tasks they rely on earlier learned tasks. The model is general instead of specific.Vickers (2017)

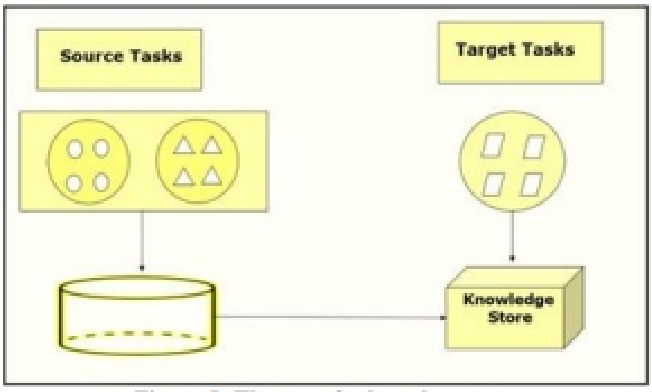


Figure 5: The transfer learning setup

In this project we have used 3 pre-trained model VGG- 16, MobileNetV2, and densenet in order to train the model and classify the results. In spite of training all the layers of the model from scratch, we have manipulated the last layer of the model as per our problem statement (to classify whether the person has wore a mask or not).

5.4.2 VGG-16 | CNN Model

VGG-16 is 16 layers deep model. The arrangement of layers in VGG-16 has been done in such a manner that it provides all the benefits of the transfer learning by overcoming the limitations of traditional machine learning approach. instead of having a large number of hyper-parameter they focused on having convolution layers of 3x3 filter with a stride 1 because there are now three ReLU units instead of just one, the decision function is more discriminative and always used same padding in order to conserve the size and maxpool layer of 2x2 filter of stride 2. [14]

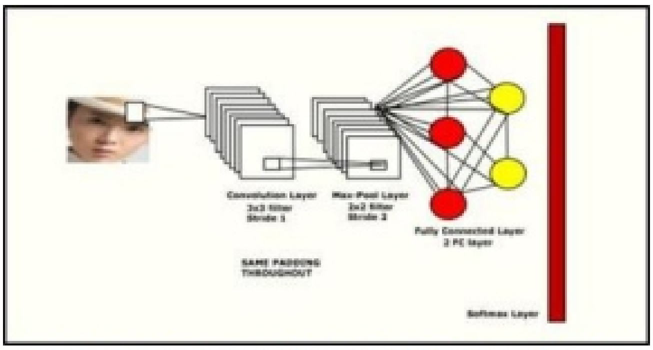


Figure 6: Arrangement of layers in VGG-16 Architecture

It follows this arrangement of convolution and max pool layers consistently throughout the whole architecture. In the end it has 2 FC(fully connected layers) followed by a softmax for output. Any manipulations that has to be done as per the problem statement can be done by altering the softmax layer. [13]

Implementation in the Project

In this project the last layer i.e softmax layer is manipulated rest all the layers have been freezed. No of epochs that has been taken is 5.

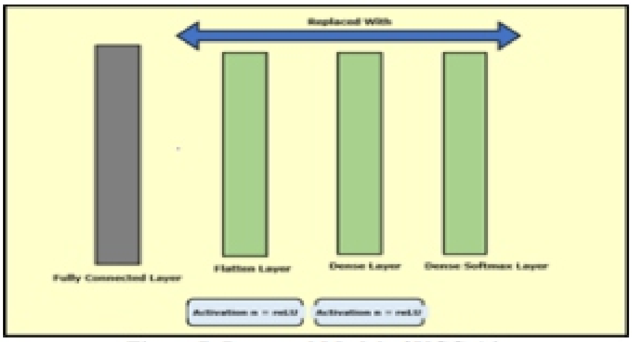


Figure 7: Proposed Model of VGG-16

5.4.3. MobileNetV2 | CNN Model

Mobilenet is a less expensive depth wise seperable conovlutional operation comprising the depth wise separable concern and point wise separable convolution operation. For classification prediction they repeat this block as 13 times followed by the usual pooling layers, fully connected layer and softmax layer.The improvement of the mobilenet is that resulted to the development of mobilenet v2 the initial layers and sums up to the output in the final layer. Second change is introduction of expansion layer followed by depth wiseseparable concern and followed by point wise. They repeat this block 17 times, pass this input to 17 layers and ends up with pooling layer followed by fully connected layer and softmax layer. This block is also called bottleneck block which reduces the computational costs.

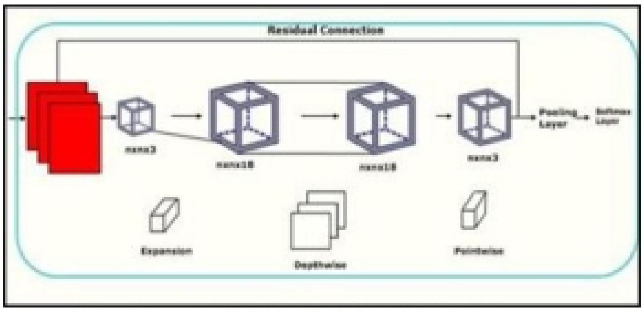


Figure 8: MobileNetV2 CNN Model Architecture

It learns the immature parameters in a more comprehensive way while performing the step of expansion in non-residual portion of the mobilenetv2. It preserves the size of the image while applying padding. Since mobilenet is specifically designed for mobile devices, it reduces the overhead of memory by decreasing the size in the pointwise seperable concern. The size of activations that need to be passed is very small from layer to layer.[18].

Implementation in the Project

In our project, the Dropout layer was used to ignore our model being over fitted with the dataset. Using MobileNetV2 (include top=False), we were able to get rid of the base layer. The pictures have been resized. The average pooling operation is used with a pool size of 128 hidden layers in our trainable model (7,7).

In the secret layer, the Relu activation function is used, and in the entire linked layer, the SoftMax activation function is used. For better accuracy, we set a learning rate of 0.01.

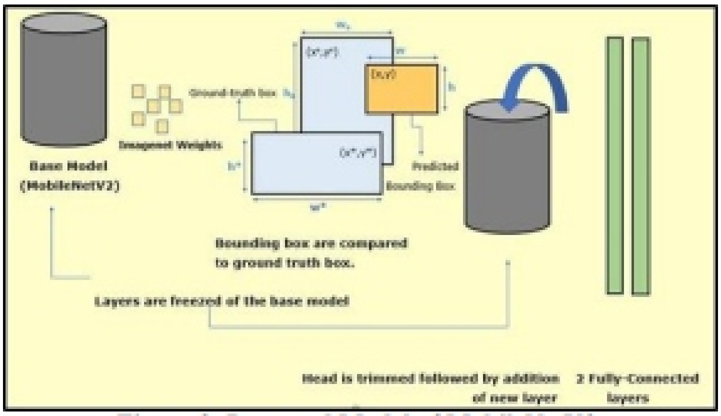


Figure 9: Proposed Model of MobileNetV2

5.4.4 InceptionV3 | CNN Model

InceptionV3 is of 48 layer network and is trained on imagenet dataset for 1000 classes. It accepts the input size of 299*299*3, a colored image of this dimension When we pass them to the first block that consists of 3 convolutional layers in which the first convolutional layer is of filter size 32 with stride 2 and a size 3, second convolu- tional layer is of filter size 32 with stride 1 and a size 3 and a second convolutional layer is of filter size 64 with stride 1 and a size 3 and we get the image size of 147x147x64. Next the max-pooling is applied, with filter size of 3x3 with stride 2 and we get a image size of 73x73*64. Followed by the max-pooling layer 2 conv layer and one max-pool layer has been used in the architecture. After that inception block A has been used 3 times followed by reduction block and again followed by inception B block 4 times followed by reduction block and inception c block 2 times. At last it is connected by 2 fully connected dense classification layers which consists of 1000 neurons because it is trained on 1000 different classes.

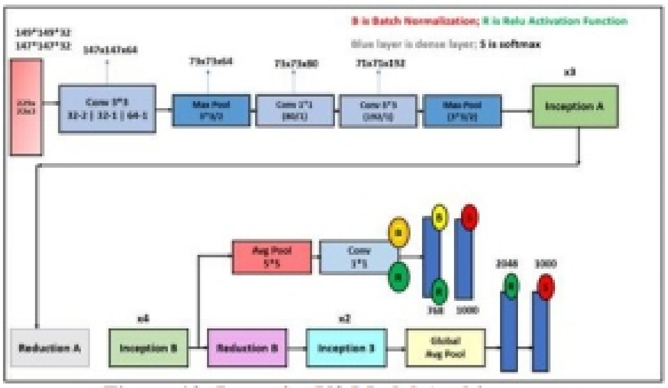


Figure 10: InceptionV3 Model Architecture

In middle of the layers again the classifiers layers has been used, in this architecture before the softmax layer auxiliaryclassifier has been used to tackle the problem of vanishing gradient. In between the layers of auxiliary classifiers are added to calculate the loss, and is to be added to final classification layer, to improve the performance. In the earlier layers the 1x1 convolution layers are added in order to reduce the dimensions, which will ultimately results into less computation resources, therefore the parameters are reduced.

Inception Block A

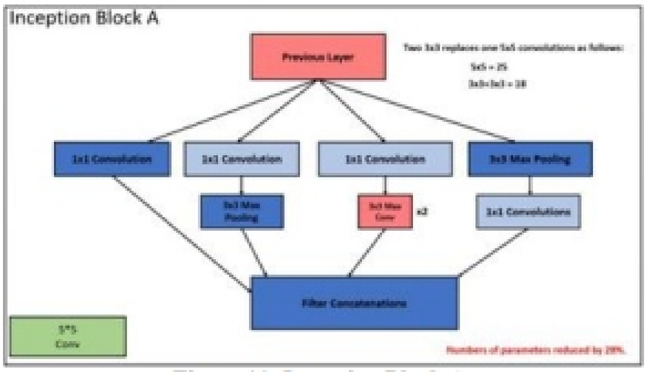


Figure 11: Inception Block A

Factorization into smaller convolutions:

In inception Block A, earlier in place of 2 layers of convolution layers(filter size of 3), one 5X5 layers of convolution has been usedHe, Zhang, Ren and Sun (2016). In earlier we were getting 25 parameters by using 1 5x5 filter (5x5 = 25), whereas now we are getting 18 parameters, by using two 3x3 filters (3x3 + 3x3 = 18). In this way the no of parameters are reduced by 28%. At the end concatenating all the layers and finding the average.

Inception Block B

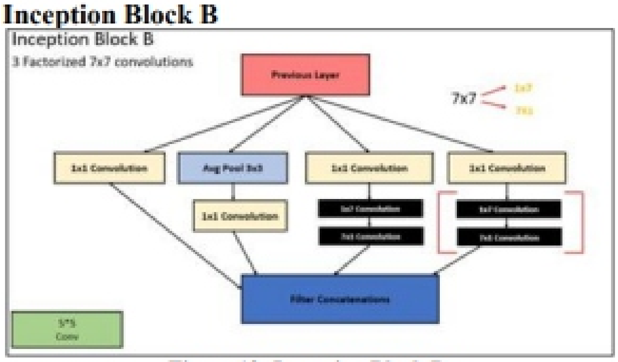


Figure 12: Inception Block B

Factorized 7x7 convolutions:

In Block B, in place of 7x7 convolutions that have been used earlier are replaced[12] by 2(7x1 + 1x7) convolution and thus the parameters are reduced.

Inception Block C

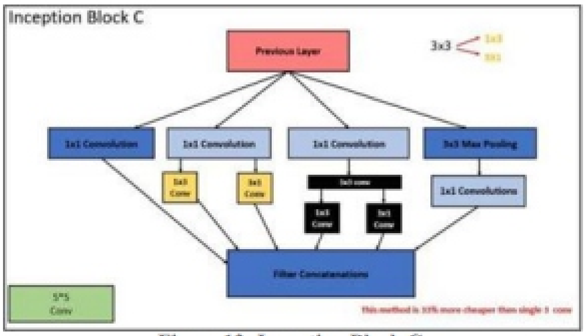


Figure 13: Inception Block C

Factorization into Asymmetric Convolutions

In inception block C, the no of parameters are again reduced by 33%. Earlier in block A, there were 3x3 filterswere used which was using 1 layer of 3x3 filter, with parameter 3x3 = 9Jogin, [12] in place of that 2 asymmetric layer of 1x3 and 3x1 filters have been used, which will result in 3x1+1x3=6 parameters. This reduces the parameters by 33%.

Reduction Block A and B

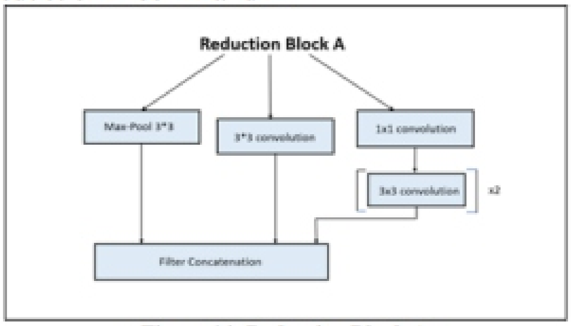


Figure 14: Reduction Block A

In reduction block A, the 5x5 convolution is reduced to 2(3x3) convolution to reduce the no of parameters and finally we are concatenating the results. [16]

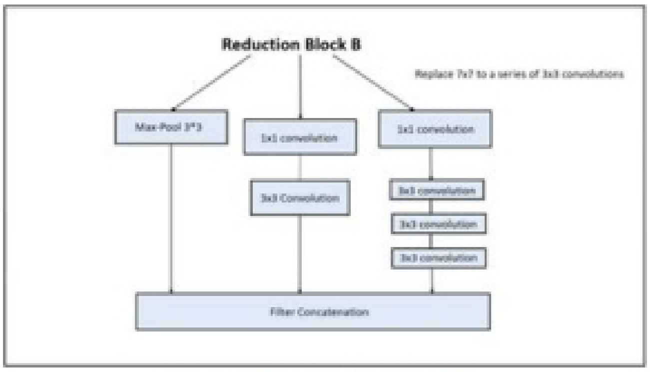


Figure 15: Reduction Block B

In reduction block B, asymmetric factorization has been performed, by replacing 7x7 convolution into (1x7 and 7x1),and thereby replacing a series of 7x7 convolution into 3x3 convolution.[4]

Implementation in the Project

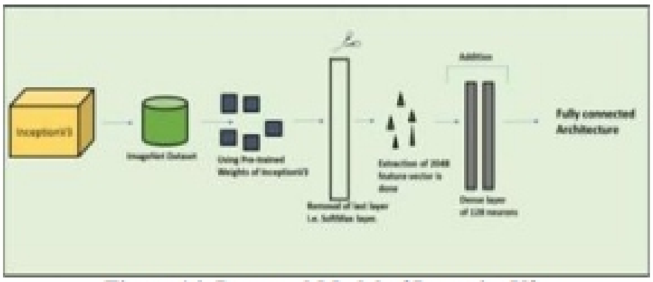


Figure 16: Proposed Model of InceptionV3

In this project pre-trained weights of InceptionV3 is used that is already trained on Imagenet dataset and the last layer is being removed i.e softmax layer and extraction of 2048 feature vector is done. After extraction of the feature vector 2 more dense layers are added of 128 neurons to build fully connected architecture to classify the image into masked and non-masked images and lastly trained on mask dataset.

5.5 Performance Evaluation Metrices

This experiment of evaluation of transfer learning mod- els for face mask detection is implemented on Google Co- laboratory (Colab Notebook) that runs on the cloud. The proposed methodology was implemented using Python and TensorFlow, the model training and tests are performed on a TESLA K80 GPU by NVIDIA[17] The transfer learning models used in this experiment are VGG16, InceptionV3, MobileNetV2 and the nature of classification results of the testing data is divided into the following four categories used for performance calculation:

True Positive (TP): Predicted that face mask is present and the original result is the same.

False Positive (FP): Predicted that mask is present but original result is no mask present.

True Negative (TN): Predicted that face mask is not present and the original result is the same.

False Negative (FN): Predicted that mask is not present but original result is mask is present.

AccuracyAccuracy is the overall number of the correct predictions fractionated by the whole number of predictions created for a dataset. It can inform us immediately if a model is trained correctly and by which method it may perform in general[14]. Nevertheless, it does not give detailed information concerning its application to the issue.
Accuracy =TP+TN/(TP+FP)+ (TN+FN)…...........(1)

Precision Precision, called PPV, is a satisfactory measure to determination, whereas the false positives cost is high[16].
Precision =TP/(TP+FP)………………………….(2)

Recall Recall is the model metric used to select the best model when there is an elevated cost linked with false negative. Recall helps while the false negatives' cost is high.[19]
Recall =TP/(TP+FN)……………………………. (3)

F1-ScoreF1-score is required when you desire to seek symmetry between both precision and recall. It is a general measure of the accuracy of the model. It combines precision and recall. A good F1-score is explained by having low false positives and also low false negatives. [25]
F1-Score =2*precission*recall/precission +recall.......

5.5.1. Confusion Matrix

It is a square matrix of size „p, used to evaluate the efficiency of a classification model (where p = number of target categories). It compares the factual goal values with the model's predictions. Here p=2 as target groups are with and without masks. Following figures shows the confusion matrices obtained for each model evaluated.[29]

VGG-16

This section describes the results for the 2 class i.e Mask and No-mask when trained on VGG-16 neural network models.

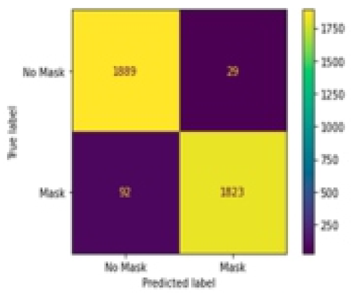


Figure 17: Confusion Matrix-VGG-16

Internal	Validation On 3833 Images	-VGG-16
Class	**Mask**	**Without Mask**
Mask	TP=1823	FP =92
Without Mask	FN=29	TN=1889
Accuracy= 0.96 **Precision**=0.98 **Recall**=0.95 **F1-Score**=0.96		

Table 1 Performance evaluation : VGG-16 Model

VGG-16 showed the accuracy of 96%, precision 98%, recall 95% followed by F1-Score of 96%.

MobileNetV2

This section describes the results for the 2 class i.e Mask and No-mask when trained on MobileNetV2 neural network models.

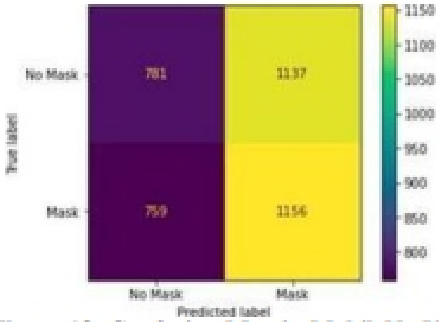


Figure 18: Confusion Matrix-MobileNetV2

Internal	Validation On 3833 Images	-VGG-16
Class	**Mask**	**Without Mask**
Mask	TP=1823	FP =92
Without Mask	FN=29	TN=1889
Accuracy= 0.96 **Precision**=0.98 **Recall**=0.95 **F1-Score**=0.96		

Table 2 Performance evaluation : MobileNetV2 Model

MobileNetV2 showed the accuracy of 50%, precision 50%, recall 60% followed by F1-Score of 54%.

InceptionV3

This section describes the results for the 2 class i.e Mask No-mask when trained on InceptionV3 neural network models.

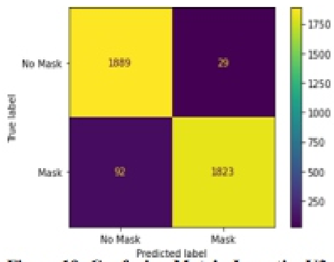


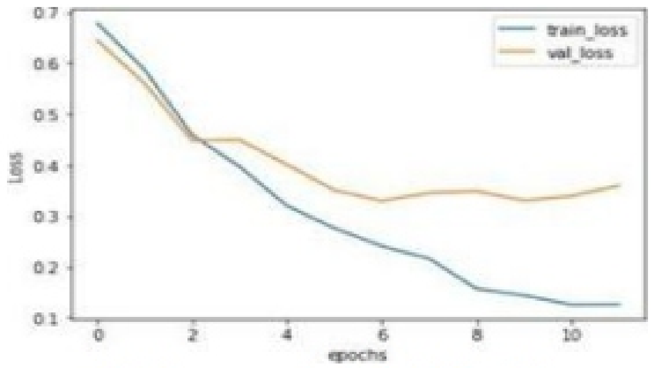
Figure 19: Confusion Matrix-InceptionV3

Internal	Validation On 3833 Images-	InceptionV3
Class	**Mask**	**Without Mask**
Mask	TP=1823	FP =92
Without Mask	FN=29	TN=1889
Accuracy= 0.96 **Precision**=0.98 **Recall**=0.95 **F1-Score**=0.96		

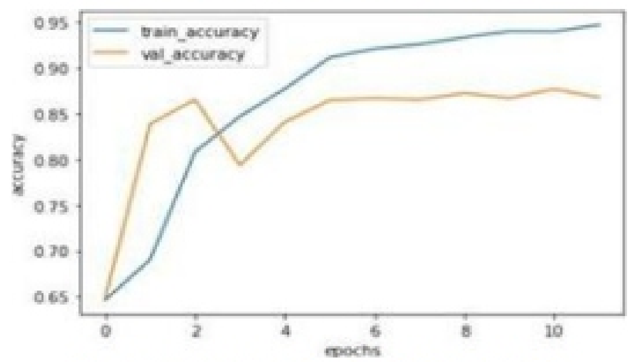
Table 3 Performance evaluation : InceptionV3 Model

InceptionV3 showed the accuracy of 98%, precision 97%, recall 99% followed by F1-Score of 98%.

5.5.2. Curve of accuracy vs epoch - VGG-16



(a) Loss vs. epochs plot for VGG model



(b) Accuracy vs. epochs plot for VGG model.
Figure 20:

Figure 20 (a) and (b) discusses accuracy vs. epoch and loss vs. epoch curves of the VGG model. VGG16 is a 16- layer deep architecture that requires a considerable number of weight parameters. As a result, it increases the inference time. As shown in Figure 23(b), it can be noted that accuracy of the VGG-Model is increasing with the increasing epochs. It learns even fine noise details from the training set. Consequently, increases the difference between training and validation accuracy. To get more insights about the performance of VGG16, a classification report is plotted, as shown in Figure 21.

	precision	recall	f1-score	support
0	0.95	0.98	0.97	1918
1	0.98	0.95	0.97	1915
accuracy			0.97	3833
macro avg	0.97	0.97	0.97	3833
weighted avg	0.97	0.97	0.97	3833

Figure 21: Classification Report - VGG-16

As mentioned in Figure 21, it can be interpreted that VGG16 classifies the images with no mask correctly with an accuracy of 95%, and images with the mask are classified correctly for about 98% time. From the above matrices, it infers that VGG16 does not generalize the data effectively on unseen data or target data. It can be interpreted from Figure
16. Around 50 images with masks have been incorrectly classified as images without masks, and 20 images of the category without masks have been classified into the category of ―with mask.‖ Further, the test set accuracy of VGG16 can be increased by adding some regularization.

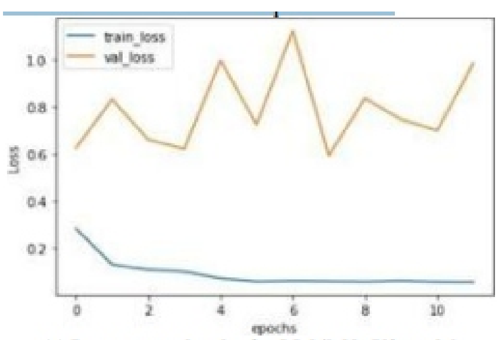
MobileNetV2

To be used at edge devices the architecture needs to be smaller and faster. MobileNetV2 chose to factorize a convolution into two steps.

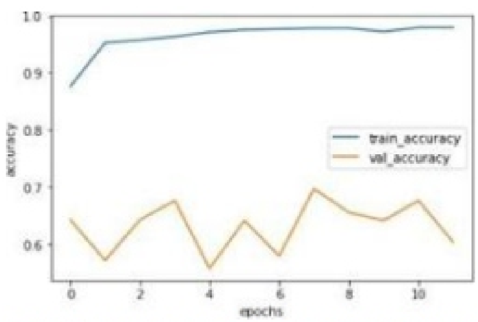
	precision	recall	f1-score	support
0	0.51	0.41	0.45	1918
1	0.50	0.60	0.55	1915
accuracy			0.51	3833
macro avg	0.51	0.51	0.50	3833
weighted avg	0.51	0.51	0.50	3833

Figure 22: Classification Report - MobileNetV2

Depth-wise separable convolutions. MobileNetV2 could reduce computation and the model size dramatically. Hence MobileNetV2 is also analyzed. The results of MobileNetV2 can be seen below. Figure 22 gives average weighted F1-score metrics as 0.51, which is considered to be the lowest among all evaluated architectures in this field of study. The reason being that the resulting network, when modified as per binary classification, gets very complex. The model learns effectively on the training data. It learns even fine details of images so accurately that it ends up miss-classifying the predicted data. As a result, the model gets saturated while evaluating the validation set of data. Below mentioned Figures 23(a) and 25(b) , shows Accuracy vs. epoch and Loss vs. epoch curves. It shows that the model learns effectively but fails to generalize effectively on test data. A primary reason being the over-learning of the trained parameters in the training phase of the model. The above confusion matrix shows that the MobileNetV2 model is giving a considerable amount of false predictions.



(a) Loss vs. epochs plot for MobileNetV2 model



(b) Accuracy vs. epochs plot for Mo- bileNetV2 model

Figure 23:

The total number of images of the category with the mask has been classified into without mask are 7 in number, while pictures of the category without the mask that are classified into with mask are total of 590 in number. It shows that MobileNetV2 predicts images with an error rate of 43.38%. The idea behind carrying out this analysis is to find the threshold resolution of images for predictions. The carried analysis is shown in confusion matrix.

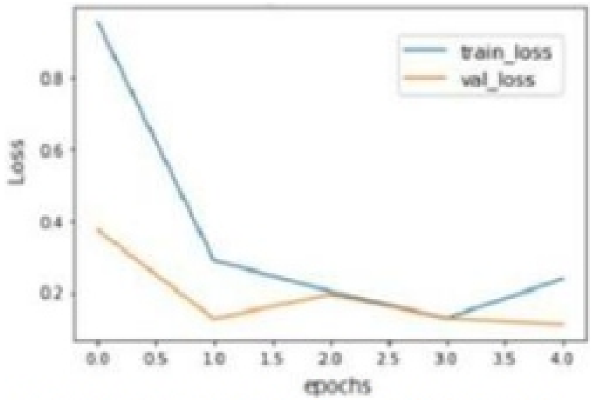
InceptionV3

Inception v3 is an image recognition model. It is made up of symmetric and asymmetric building blocks, including convolutions, average pooling, max pooling, concatenations, dropouts, and fully connected layers. Batch normalization is used extensively throughout the model and applied to activation inputs. Loss is computed using Softmax.

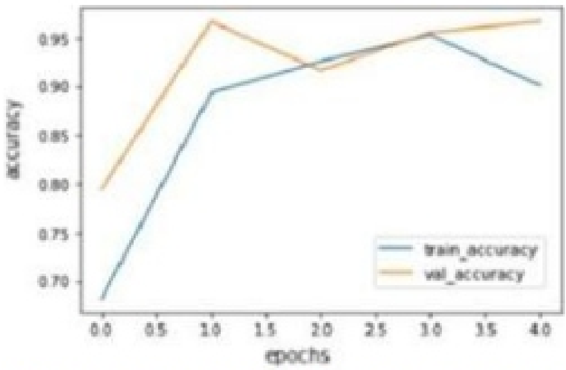
	precision	recall	f1-score	support
0	0.95	0.98	0.97	1918
1	0.98	0.95	0.97	1915
accuracy			0.97	3833
macro avg	0.97	0.97	0.97	3833
weighted avg	0.97	0.97	0.97	3833

Figure 24: Classification Report - InceptionV3

Figure 24 gives average weighted F1-score metrics as 0.98, which is considered to be the highest among all evaluated architectures in this field of study. The reason being that it optimizes the network for better model adaptation and computationally being less expensive.



(a) Loss vs. epochs plot for InceptionV3 model



(b) Accuracy vs. epochs plot for InceptionV3 model

Figure 25:

Above mentioned Figures 25(a) and (b) , shows Accu- ar cy vs. epoch and Loss vs. epoch curves. It shows that the model learns effectively and able to generalize effectively on test data.

VI. WORKING OF THE SYSTEM

After evaluating the proposed face mask detection moels, in this step, the best model with high accuracy rate (MobileNetV2, VGG-16, and inceptionV3) will be applied to the embedded vision system.
When someone is not wearing a face mask, it will be designated with a red box around their face with the text, ―No Mask,‖ and when wearing a face mask, it will be seen a green box around their face with the text, ―Mask‖ The same thing is depicted by the above figure. On the other hand, the proposed model

with social distancing task detects peoples and provides the bounding box information. After that, the Euclidean distance between each detected centroid pair iscomputed using the detected bounding box and its centroid information based on (x, y) dimensions for each bounding box.

VII. COMPARISON OF THE RESULTS

System consists of face mask dataset which comprises of the images of masked and non-masked face. Images are pre-processed and given for the transfer learning models for training and it classifies whether person is wearing a face mask or not. By using various parameters such as accuracy, precision, recall and f1-score we can find how good our model is trained and gives the results.

Comparison of 3 transfer learning model VGG-16, MobileNetV2 and InceptionV3 has been done for the face mask detection.

Same dataset has been used for training and testing.

80% of the images has been used for training and 2% for testing

VGG-16 has got an accuracy of 91%.

MobileNetV2 has got an accuracy for 50%.

InceptionV3 has got an accuracy of 98%.

Compared to VGG-16 and MobileNetV2, the InceptionV3 got the best accuracy

Method	Classification and Detection	Performance Metrics Used	Performance Recorded
VGG-16	Yes	Accuracy, Precision, Recall, F1-Score.	Accuracy was 91% Precision was 93% Recall was 89% F1-Score was 90%
MobileNetV2	Yes	Accuracy, Precision, Recall, F1-Score.	Accuracy was 50% Precision was 50% Recall was 60% F1-Score was 54%
InceptionV3	Yes	Accuracy, Precision, Recall, F1-Score.	Accuracy was 98% Precision was 97% Recall was 99% F1-Score was 98%

Table 4 Comparison of the results

VIII. CONCLUSION

The epidemic caused by the COVID-19 disease has caused an upsurge of infected cases worldwide. The governments of several nations around the world have enforced obligatory use of protective face masks to prevent the trans- mission of the virus and prevent infections. But manual monitoring of public for use of face masks is a tedious process. Hence the development of automated face mask detector has been of the ultimate need of the society in this pandemic phase In this paper, we have proposed three models of face mask detection using pre-trained models VGG16, MobileNetV2 and InceptionV3 for detecting face mask and evaluated their performance using various metrics. The results showed that InceptionV3 models have produced outstanding accuracies on the given dataset. Our future work will concentrated on the development of automated face mask detector using hybrid model.

ACKNOWLEDGEMENT

I am grateful to my guide Dr. VaishaliJadhav with whom I had the pleasure to work in this project. She had provided me the extensive personal and professional guidance and taught me a great deal about both scientific research and life in general.

DECLARATIONS

The Author declare that they have no known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCE

[1] Altmann, D.M., Douek, D.C., Boyton, R.J., 2020. What policy makers need to know about covid-19 protective immunity. The Lancet 395, 1527– 1529.

[2] Chen, Y., Hu, M., Hua, C., Zhai, G., Zhang, J., Li, Q., Yang, S.X., 2021. Face mask assistant: Detection of face mask service stage based on mobile phone. IEEE Sensors Journal 21, 11084–11093.

[3] COVID, W., 2020. Dashboard–egypt: https://covid19. who. int/region/emro/country/eg. Accessed on 29.

[4] Dietterich, T.G., et al., 2002. Ensemble learning. The handbook of brain theory and neural networks 2, 110–125.

[5] Din, N.U., Javed, K., Bae, S., Yi, J., 2020. A novel gan-based network for unmasking of masked face. IEEE Access 8, 44276–44287.

[6] Ejaz, M.S., Islam, M.R., Sifatullah, M., Sarker, A., 2019. Implementation of principal component analysis on masked and non-masked face recog- nition, in: 2019 1st international conference on advances in science, engineering and robotics technology (ICASERT), IEEE. pp. 1–5.

[7] Feng, S., Shen, C., Xia, N., Song, W., Fan, M., Cowling, B.J., 2020. Rational use of face masks in the covid-19 pandemic. The Lancet Respiratory Medicine 8, 434–436.

[8] Fouquet, H., 2021. Paris tests face-mask recognition software on metro riders. Bloomberg. https://www. bloomberg. Com/news/articles/2020-05-07/paris-tests-face-mask recognition-software-on-metro-riders. Accessed 31.

[9] Gazzah, S., Bencharef, O., 2020. A survey on how computer vision can response to urgent need to contribute in covid-19 pandemics, in: 2020 International Conference on Intelligent Systems and Computer Vision (ISCV), IEEE. pp. 1–5.

[10] He, K., Zhang, X., Ren, S., Sun, J., 2016. Proceedings of the ieee conference on computer vision and pattern recognition .

[11] Hussain, S.A., Al Balushi, A.S.A., 2020. A real time face emotion classification and recognition using deep learning model, in: Journal of Physics: Conference Series, IOP Publishing. p. 012087.

[12] Jogin, M., Madhulika, M., Divya, G., Meghana, R., Apoorva, S., et al., 2018. Feature extraction using convolution neural networks (cnn) and deep learning, in: 2018 3rd IEEE international conference on recent trends in electronics, information & communication technology (RTE- ICT), IEEE. pp. 2319–2323.

[13] Kaur, G., Sinha, R., Tiwari, P.K., Yadav, S.K., Pandey, P., Raj, R., Vashisth, A., Rakhra, M., 2021. Face mask recognition system using cnn model. Neuroscience Informatics , 100035.

[14] Kawulok, M., Celebi, E., Smolka, B., 2016. Advances in face detection and facial image analysis. Springer.

[15] Khan, M.K.J., Ud Din, N., Bae, S., Yi, J., 2019. Interactive removal of microphone object in facial images. Electronics 8, 1115.

[16] Kleinbaum, D.G., Dietz, K., Gail, M., Klein, M., Klein, M., 2002. Logistic regression. Springer.

[17] Larsen, D., 2020. Homemade cloth face masks to fight the covid19 pandemic; a call for mass public masking with homemade cloth masks . Li, C., Wang, R., Li, J., Fei, L., 2020. Face detection based on yolov3, in: Recent Trends in Intelligent Computing, Communication and Devices. Springer, pp. 277–284.

[18] Loey, M., Manogaran, G., Taha, M.H.N., Khalifa, N.E.M., 2021. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the covid-19 pandemic. Measurement 167, 108288.

[19] Loey, M., Smarandache, F., M Khalifa, N.E., 2020. Within the lack of chest covid-19 x-ray dataset: a novel detection model based on gan and deep transfer learning. Symmetry 12, 651.

[20] Mangayarkarasi, T., Anusha, S.S., Pranavikha, P., 2021. Covid-19 mask detection using matlab based image processing. Ilkogretim Online 20, 5016–5023.

[21] Naseem, I., Togneri, R., Bennamoun, M., 2010. Linear regression for face recognition. IEEE transactions on pattern analysis and machine intelligence 32, 2106–2112.

[22] Navada, A., Ansari, A.N., Patil, S., Sonkamble, B.A., 2011. Overview of use of decision tree algorithms in machine learning, in: 2011 IEEE control and system graduate research colloquium, IEEE. pp. 37–42.

[23] Nieto-Rodriguez, A., Mucientes, M., Brea, V.M., 2015. System for medical mask detection in the operating room through facial attributes, in: Iberian Conference on Pattern Recognition and Image Analysis, Springer. pp. 138–145.

[24] Park, J.S., Oh, Y.H., Ahn, S.C., Lee, S.W., 2005. Glasses removal from facial image using recursive error compensation. IEEE transactions on pattern analysis and machine intelligence 27, 805–811.

[25] Qin, B., Li, D., 2020. Identifying facemask-wearing condition using image super-resolution with classification network to prevent covid-19.Sensors 20, 5236.

[26] Sandesara, A.G., Joshi, D.D., Joshi, S.D., 2020. Facial mask detection using stacked cnn model. Int. J. Sci. Res. Comput. Sci. Eng. Inform. Technol .Terzi, D.S., Terzi, R., Sagiroglu, S., 2017. Big data analytics for network anomaly detection from netflow data, in: 2017 International Conference on Computer Science and Engineering (UBMK), IEEE. pp. 592–597.

[27] Ting, D., 2020. W., carin l., dzau v., wong ty digital technology and covid- 19. Nature Med 26, 459–461.

[28] Tu, P.L., Chung, J.Y., 1992. A new decision-tree classification algorithm for machine learning, in: TAI'92-Proceedings Fourth International Confer- ence on Tools with Artificial Intelligence, IEEE Computer Society. pp. 370–371.

[29] Vickers, N.J., 2017. Animal communication: when i'm calling you, will you answer too? Current biology 27, R713–R715.

[30] Wang, Z., Wang, G., Huang, B., Xiong, Z., Hong, Q., Wu, H., Yi, P., Jiang, K., Wang, N., Pei, Y., et al., 2020. Masked face recognition dataset and application. arXiv preprint arXiv:2003.09093 .

[31] Wen, L., Li, X., Gao, L., 2020. A transfer convolutional neural network for fault diagnosis based on resnet-50. Neural Computing and Applications 32, 6111–6124.

ISSN :25060

CONTENT-AWARE VIDEO ANALYSIS OF SPORT VIDEOS USING MACHINE LEARNING

[1] SEJAL AHUJA, [2] PARIKSHITA NALWADE,3KIRTI KADAM,4HARSHADA RASKAR,5DEEPTI PANDE

1,2,3,4,5 Department of Information Technology, Trinity College of Engineering and Research, Pune, India

E-mail: 1sejalahuja77@gmail.com

A B S T R A C T

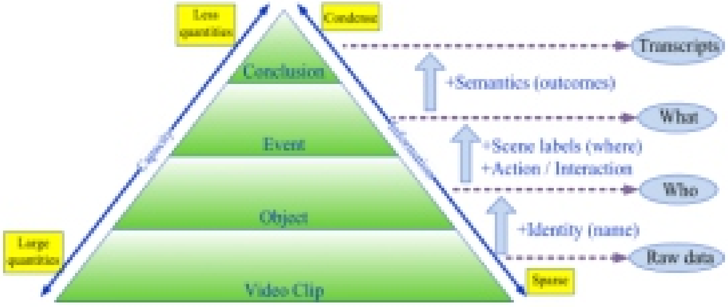
The games content examination has expanded quickly in ongoing many years, in view of the such tremendous development of video transmission over the Web and the interest for computerized telecom applications. In which enormous information produced by sport recordings so there is testing errand of mining or characterization of game video information spreading over web. Existing studies have zeroed in on the techniques of sports video examination from the spatiotemporal perspective (Meta information) rather than a substance based viewpoint(actions). Rationale behind proposed work is to parse sport recordings in light of activities only which sort of game is playing in current video grouping rather than crude data about that video. The game video broadcast is the vitally satisfied spreading and requesting over web. The monstrous interest for sports video broadcasting, many game applications, for example, hot-star, you-tube and numerous webbased entertainment applications. Ongoing many years the games programs have turned into a prevailing concentration in the field of diversion. Research on big data analytics has attracted much attention to machine learning and artificial intelligence techniques. Thus, there is need major areas of strength for of which can deal with such colossal game information in light of content of videos.

Keywords: Action Recognition, Content Aware System, Content-based Multimedia Analysis, Event Detection, Semantic Analysis, Sports, Survey.

I. INTRODUCTION

Presently a days bunches of web contents created consistently by social and game applications, for example, hot-star, begin sport channels, jio-tv, youtube and so on. In this way, there is additionally need of sports video examination frameworks which ought to be group video contents proficiently. The video content mining depends on AI approaches in light of their profound learning degree. The overall technique of a substance mindful video investigation framework incorporates highlight extraction, data thinking, and information game plan. The video items can be ordered by utilizing meta information of recordings too relevant highlights of recordings. The proposed work zeroed in on satisfied content aware video characterization rather than spatiotemporal perspective. Sports information examination is turning out to be progressively enormous scope, broadened, and shared, yet trouble perseveres in quickly getting to the most significant data.

Past reviews have zeroed in on the strategies of sports video examination from the spatiotemporal perspective rather than a substance based perspective, and not many of these investigations have thought about semantics. This study fosters a more profound translation of content aware games video examination by inspecting the knowledge presented by investigation into the construction of content under various situations. Based on this knowledge, we give an outline of the subjects especially pertinent to the examination on happy mindful frameworks for broadcast sports. Specifically, we focus on the video content This paper presents an exploration of the evolution of sports video analysis over the last 10 years. The focus is on content-aware analysis methods, including object-, event-, and contextoriented groups, applied in broadcast sportscasts. Each gathering focuses on specific topics related to the subject at hand and provides insights into current trends and challenges that form a framework for future developments in this specialized area. We advance our findings for sports video analysis by investigating three prominent models: a substance model, a progressive model, and a pattern/difficulty model Content Pyramid: a hierarchical content model.



II. LITERATURE REVIEW

As per the construction of the substance pyramid, we studied best in class strategies from the part of semantic level, to be specific, feature location and occasion acknowledgment, object discovery and action recognition, and contextual inference and semantic analysis.

[1]. Nian Liu , 1 Lu Liu,2 and Zengjun Sun 2, Research Article "Football Game Video Analysis Method with Deep Learning": In this system, a deep learning method is used to design a football event detection algorithm for football game video analysis. The algorithm can naturally identify and group different occasions in football game videos. Among them, the three-layered convolutional network is utilized for highlight extraction, which can deal with various edges of pictures simultaneously, to hold pertinent data between outlines. It utilizes a bidirectional intermittent organization to incorporate elements from both positive and negative bearings to get past and future relevant data to work on the impact of occasion discovery. primary substance is partitioned into two sections: (1) characterization of football occasions. grouping model utilizes the 3D CNN organization and the SoftMax classifier for include extraction and prescient order for event fragments, separately. As indicated by the qualities of the football game video, the model info is separated into a fulloutline picture and a focal region of the edge, which are individually placed into 3D CNN to remove elements, and component combination is performed. SoftMax classifier computes the anticipated worth of every classification for the occasion portion and chooses the one with the biggest anticipated esteem as the anticipated classification of the event. (2) Football event detection: The event discovery model depends on the grouping model by adding BLSTM design to all more likely get dynamic data between different casing arrangements. During preparing, the dataset is extended first and afterward it is extended utilizing SGD calculation. During testing, a sliding window is utilized to fragment each part into frame segmentation algorithms together with video expansion and segmentation and then opts for segments that include more than 30 cells per square foot. Through filtering and merging, the beginning and end limits of events are additionally affirmed, and category labels are created. To check the legitimacy and rightness of proposed technique, complete and orderly investigations are done, and the model is dissected from various viewpoints. exploratory outcomes affirm the possibility of this work.

III. RESEARCH METHODOLOGY

This study aims to overcome existing drawbacks of video classification in sport videos.

Our work is based on machine learning techniques for video analysis based on content or actions in videos instead of meta data of videos. We are going to develop video analysis module wise: 1) Sport Video Input 2) Frame extraction 3) Feature extraction 4) Video object detection 5) Video object recognition 6) Action recognition 7) Video classification

K-Convex hull Algorithm: The well-known understanding about convex hull is that it is Minimal perimeter problem for sets containing a fixed set and convex hull Co(E) of E is the bounded connected set constituting a minimization problem. Convex hull algorithms are broadly divided into two categories. 1. Graph traversal 2. Incremental The graph traversal algorithms construct CHs by identifying some initial vertices of CH and later finding the remaining points and edges by traversing it in some order. The Gift Wrapping, Graham scan, and Monotone chain are such algorithms. Incremental algorithms first find an initial CH and then insert or merge the remaining points, edges or even sub-CHs as they are discovered, into current CH sequentially or recursively to obtain the final CH.

CNN Algorithm: Image processing using Convolutional neuronal networks (CNN) has been successfully used in various fields of action, for example, geo procedures, structural designing, mechanics, modern observation, insubordination division, automatics and transport. Image preprocessing, date decrease, division and acknowledgment are the cycles utilized in overseeing pictures with CNN. Each info neuron addresses variety data in the picture, and each result neuron relates to a picture. All pictures will be scaled to a similar size (width and level) and little to rush to learn. On the measures of the pictures not entirely settled on the size of the information vector and the quantity of neurons. The exchange capability for this kind of issue is called sigmoid capability. The pace of learning has values in the reach [0.1] and the blunder it is prescribed to have under 0.1.

Processing of pictures with CNN includes various cycles, for example,

1. Image pre-handling, an activity which shows an image (contrast upgrade, sound decrease) with similar aspects as the first picture. The goal of pictures prehandling with CNN comprises in improving, re establishing or revamping pictures. The settled issues are the cartographic sorts, to upgrade a capability, a guess capability for the recreation of image. 2. Information decrease or feature extraction includes extricating various elements less than the quantity of pixels in the info window. The activity comprises in compacting the picture followed by separating mathematical attributes (edges, corners, joints), facial highlights, and so forth.

IV. LITERATURE SEARCH

Many different databases were used during the literature review, like IEEE, MDPI, Hindawi, Wiley digital library, etc. The document exploration was distributed using the keywords "Deep learning", "Computer Vision in Sports" and "Video Action Recognition" by the research's well-defined objectives and exploratory queries. For the review, a complete of over 100+ publications from 2015 to 2021 are retrieved.

V. RESULTS

This outcome section displays, the investigation's findings, and recommendations. Out of 43 articles, 15 papers are included in the final set of papers. The papers are taken from a variety of scholarly journals and meetings between 2015 and 2021. Hind awi is at the top of the list in this review among the journal categories based on the results. The primary goal is to verify the validity and correctness of proposed method, comprehensive and systematic experiments are carried out, and the model is analysed from different aspects. experimental results confirm the feasibility of this work.

BIBLIOGRAPHY

[1] Lu Liu,2 and Zengjun Sun 2, Research Article "Football Game Video Analysis Method with Deep Learning" Business School, Hindawi,ComputationalIntelligence and Neuroscience Volume 2022.

[2] Fei Wu, Qingzhong Wang, Jiang Bian, "Survey on Video Action Recognition in Sports: Datasets, Methods and Applications", 2022.

[3] BanothThulasya Naika, Mohammad Farukh Hashmia, Neeraj Dhanraj Bokdeb, "A Comprehensive Review of Computer Vision in Sports: Open Issues, Future Trends and Research Directions", A) National Institute of Technology, Warangal, India B) Department of Engineering - Renewable Energy and Thermodynamics, Aarhus University, 8000, Denmark.

[4] Brandon Giles, Stephanie Kovalchik & Machar Reid, "A machine learning approach for automatic detection and classification of changes of direction from player tracking data in professional tennis", journal of Sports Sciences, 2019.

[5] Huang-Chia Shih, Member, IEEE," A Survey on Content aware Video Analysis for Sports", IEEE Transactions On Circuits And Systems For Video Technology, January 2017.

[6] Xun Gong1 and Fucheng Wang 2,"Research Article Classification of Tennis Video Types Based on Machine Learning Technology", Wireless Communications and Mobile Computing Volume, 2021.

[7] Wan Ding, Dong-Yan Huang, Zhuo Chen, Xinguo Yu, Weisi Lin, "Facial Action Recognition using Very Deep Networks for Highly Imbalanced Class Distribution", Proceedings of APSIPA Annual Summit and Conference, 2017.

[8] Yang Bo, Wenbo He, "Few-Shot Learning of Video Action Recognition Only Based on Video Contents" Yangdi Lu

[9] Pierre-Etienne Martin1•Jenny Benois-Pineau1 • Renaud Peteri ´ 2 •Julien Morlier,"Fine grained sport action recognition with Twin spatio-temporal convolutional neural networks", Springer, 2020.

[10] Ahmed Ghozia,GamalAttiya,EmadAdly,and Nawal El Fishawy, "Research Article Intelligence Is beyond Learning: A Context-Aware Artificial Intelligent System for Video Understanding", Hindawi Computational Intelligence and Neuroscience Volume, 2020.

[11] Yao Lu, Aakanksha Chowdhery, Srikanth Kandula, "Optasia: A Relational Platform for Efficient Large-Scale Video Analytics", ACM, 2016.

[12] Julius Žemgulys1, Vidas Raudonis1,Rytis Maskeliūnas2,Robertas Damaševičius3, "Recognition of basketball referee signals from real-time videos", Journal of Ambient Intelligence and Humanized Computing, 2018.

[13] Behzad Mahaseni, Erma Rahayu Mohd Faizal, And Ram Gopal Raj, "Spotting Football Events Using Two-Stream Convolutional Neural Network and Dilated Recurrent Neural Network", Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, University of Malaya, Malaysi, IEEE, 2019.

[14] Genki Suzuki 1, ShoTakahash, Takahiro Ogawa And Miki Haseya, "Team Tactics Estimation In Soccer Videos Based on a Deep Extreme Learning Machine and Characteristics of the Tactics",1) Faculty of Information Science and Technology, Hokkaido University, 2) Faculty of Engineering, Hokkaido University, IEEE, 2019.

[15] Ruikang Liu 1, Qing Han 1,Weidong Min2,3,Linghua Zhou 1 and Jianqiang Xu "Vehicle Logo Recognition Based on Enhanced Matching for Small Objects, Constrained Region and SSFPD Network", School of Information Engineering, Nanchang University, MDPI, 2019.

[16] Aftab Alam,Irfan Ullah And Young-Koo Lee, "Video Big Data Analytics in the Cloud: A Reference Architecture, Survey, Opportunities, and Open Research Issues", IEEE, 2020.

[17] Weihua Xu, Md Azher Uddin ID , Batjargal Dolgorsuren, Mostafijur Rahman Akhond, Kifayat Ullah Khan, Md Ibrahim Hossain and Young-Koo Lee, "Similarity Estimation for Large-Scale Human Action Video Data on Spark", Data and Knowledge Engineering Lab, Department of Computer Science

and Engineering, Kyung Hee University, MDPI, 2018.
[18] LinhongWang,Yunhao Wang,2 and Yiming Bie,"Research Article Automatic Estimation Method for Intersection Saturation Flow Rate Based on Video Detector Data", Hindawi Journal, 2018.
[19] Hanli Wang 1, Xiaobin Zheng 2, Bo Xiao 3, "Large-Scale Human Action Recognition with Spark", Department of Computer Science and Technology, Tongji University, IEEE, 2015.
[20] Huang-Chia Shih, "A Survey of Content-Aware Video Analysis for Sports", IEEE, 2017.

ISSN :25060

DIABETIC RETINOPATHY DETECTION SYSTEM BY MACHINE LEARNING

[1]AKSHATA LANDE, [2]BHAGVATI GHOLAP,3SNEHAL SATHE,4GAURAV PHADTARE, 5DEEPTI PANDE

Department of Information Technology KJ's Educational Institute
Trinity College of Engineering and Research, Pune, India
E-mail: akshatalande@gmail.com

A B S T R A C T

Diabetic Retinopathy (DR) is an eye disease in humans with diabetes which may harm the retina of the eye and may cause total visual impairment. Therefore it is critical to detect diabetic retinopathy in the early phase to avoid blindness in humans. Our aim is to detect the presence of diabetic retinopathy by applying machine learning algorithms. Hence we try and summarize the various models and techniques used along with methodologies used by them and analyze the accuracies and results. It will give us exactness of which algorithm will be appropriate and more accurate for prediction. Machine learning consist of a number of stages to detect retinopathy in the images that includes converting image to suitable input format, various preprocessing techniques. It also includes training a model with a training set and validating with a different testing set. Method proposed in this project can be listed Image Preprocessing, and Supervised learning and Feature Extraction. First, the images are preprocessed. They are converted. Proper resizing of image is also done. As the images are heterogeneous they compressed into a suitable size and format. The main objective of this work is to build a stable and noise compatible system for detection of diabetic retinopathy.

Keywords- Retinopathy, Machine Learning Algorithms

I. INTRODUCTION

This project presented the development of an automated system for diabetic retinopathy detection in color retina images, through the implementation of machine learning techniques. Diabetic retinopathy is one of the common complications of diabetes. It's a severe and widely spread disease among diabetic. People suffering from diabetes are at high risk of developing various eye diseases over time. Traditionally, detecting DR is a time-consuming and manual process, which requires an ophthalmologist or trained clinician to examine and evaluate digital color photographs of the retina, to identify the presence of vascular abnormalitiescaused by the DR.

There are different stages of DR

- NPDR(Non-Proliferative Diabetic Retinopathy)
- MildDR
- PDR(Proliferative Diabetic Retinopathy)

NPDR is the earliest stage of DR, In these stage blood vessels in retina startsleaking fluid and some amount of blood in eye Mild DR is the mild stage has onemicro-aneurysm (MA), a small circular red dot at the end of blood vessels. PDR isthe another stage of DR, In these stage blood vessels in retina close and flow ofblood fluid stopped by vessels because of blockage pressure is build up in eyewhich can damage optic nerve PDR is more dangerous than NPDR therefore earlydetection of DR is important. People with DR whose eye sight is at risk can betreated with laser, to prevent visual blindness. But currently there is no treatmentthat can restore the vision that has already been lost therefore Early detection ofDR is important to stop further damage of eye and to save patient life.The proposed methodology is to explore machine learning technique todetect diabetic disease using thermal images of an eye and to introduce the effectof thermal variation of abnormality in the eye structure as a diagnosis imagingmodality which are useful for ophthalmologists to do the clinical diagnosis.

In the proposed Project work to design and implement a system that can beprovide eye diabetic disease detection using thermal image, the system carriedout various features extraction using image segmentation and use CNN machine learning classification algorithm to detect the actual disease.

II. LITERATURE REVIEW

NandanaPrabhu et al., [1] Have proposed a system for Diabetic Retinopathy detection based on thepresence of the feature that shows the symptoms of the disease. Thesystem makes use of fundus images, the bright lesions on the retinaand the exudates are extracted as they indicate the symptoms of the disease.Based on the features extracted various stages of the disease is detected using hierarchical classification.They have emphasized the needs of detection systemdue to increased number of cases and less ophthalmologists to treat, and thesystem has resulted in high accuracy in sensitivity and specificity.

Anupriyaa Mukherjee et al., [2]Have discussed about various image processing techniques to classify thenormal and diseased image in order to find the problems in the detection system.Preprocessing and detection of the various features such as Optic disc detection,Blood vessels extraction, and Exudates detection are applied. They have proposeda low cost system to diagnose at an early stage with a high accuracy rate; thealgorithm could be used even for poor computing system.

Imran Qureshi et al., [3]Have presented about review of CAD systems in iagnosing DiabeticRetinopathy. They have also discussed about all the CAD systems which have beendeveloped for various needs such computation Intelligence and Image processing techniques. They also conducted a survey on screening algorithmsvarious research papers in detection and their challenges and results.Demonstration the challenges and automated DR methodologies along withpossible solutions were demonstrated.

YogeshKumaran, Chandrashekar M. Patil [4]Have presented a brief survey in detecting DR using different preprocessingand segmentation techniques as it is difficult to process the raw fundus images bymachine learning algorithms. They have given brief view from the nutshell in orderto facilitate others on recent advancement and research for their work. This canalso help in the insight detection which is based on the work of researches in thefield.

Detection of Diabetic Retinopathy using Machine Learning (IRJET)"Nov 2020 Summary : This paper summarizes study of few literatures related to the detection of Diabetic Retinopathy .Explored the potential usage of the CNN in retinal image classification. Due to the manual methods by medical personnel, an automated system can reduce the labor involved in diagnosing large quantities of retinal images significantly. Content/Idea : CNN in retinal image classification 2.1.2

"A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection" IEEE October 15, 2019 Summary : In this paper, focused to classify all the stages of DR, especially the early stages, which is the major shortcoming of existing models.

And proposed a CNN ensemble-based frame work to detect and classify the DR‟s different stages in color fundus images. Content/Idea : Classification of the DR‟s in different stages .

Summary : Developed the automatic detection of blood vessel in the retinal images. This paper introduced the automatic detection of the retinal blood vessels and the measurement of the vessel diameter. In which it is important for the detect and the treatment of different ocular diseases including Diabetic Retinopathy, glaucoma and hypertension .The proposed method consists of three main steps: Pre-processing of retinal images, Vesselness filter is used to enhance the blood vessels and Finally Hessian multi-scale enhancement filter is designed from the adaptive thres holding of the output of a vesselness filter for vessels detection. Content/Idea:Automatic detection of the retinal blood vessels.

III. RESEARCH METHODOLOGY

There are various methods used in process of detection of DR. Divide the process into three steps pre-processing, the feature of DRdetection and classification.

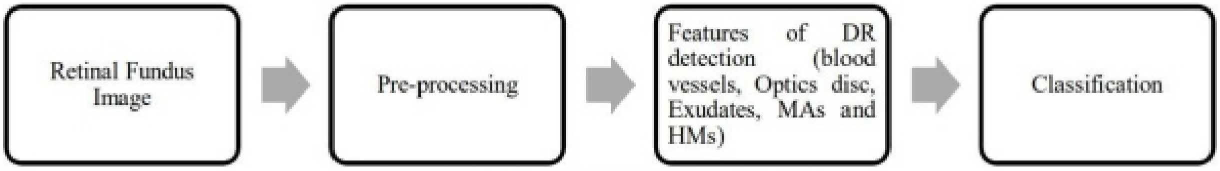


Fig. 3: flow chart of Diabetic retinopathy [10]

A. Pre-processing

Pre-processing is an important step in process of detecting DR. Most of the researchers convert retinal fundus image into green channel [1-3], [5-8], [10-14] as a contrast of MAs and HMs are high in this channel. Contrast limited adaptive histogram equalization (CLAHE) is applied to enhance the image and then median filter to remove noise [4, 6, 7 and 10]. 2D wavelet transform and median is also applied as pre-processing in [4]. Resize the image to reduce processing time [4, 6, 7 and 10]. Gaborfilter is used to enhance blood vessels [5]

B. The Feature of DR detection

1) Blood Vessels detection: MAs, HMs, and EXs are the main reason for blood vessel damage. So the blood vessels detection is animportant factor for detection of DR. Blood vessels area in the normal retinal image is approximately 37230.57 [7] but DR imagehas less than this area because of contraction.The Morphological operation followed by thresholding is applied to detect [2, 12]blood vessels. Retinal fundus image is converted into HSV color image then applied bi-cubic interpolation, gamma correction,median filter, and thresholding technique to segment blood vessels [6]. Match filter and firstorder derivation of Gaussianmatched filter technique are applied in [8].

Fuzzy C-Means clustering and morphological operations are used to detect bloodvessels [10]. The Retinal fundus image is converted into a CMY color image after that isolate Magenta channel [11]. On thatchannel morphological operations and thresholding is applied.

2) Optic disc detection: The Optic disc is bright yellow portion and approximately circular in shape [6]. To detect EXs, we need todetect a portion of the optic disc in retinal fundus image because Optic disc and EXs both have same intensity [7-8]. Optic discmark was created using brightest pixels in an image and its coordinates are a center of the marks [2, 10]. Three facts are used todetect optic disc [6]. The first fact is image acquisition and prior information such as a left or right eye. The second fact is arepresentation of Optic disk as a bright yellow portion. The Third fact is optic disc is circular in shape. The Morphologicaloperation followed by watershed segmentation is used to segment optic disc in [8]. Create a binary image of Difference between agreen channel and cyan channel of retinal fundus image after that dilation and Hough transform are applied to create optic discmark [11].

3)EXs Detection: Exudates are bright yellow objects in retinal fundus image which have less than 100 pixels in fundus image [2].Maximum entropy double thresholding technique is applied to define both high and low bright features without affecting the background [4]. Finally, to detect exudates remove all the objects which are greater than 3000 pixels. A brightness of fundusimage is changed using nonlinear curve with HSV (Hue saturation value) space [6] and then gamma correction applied on a red and green component to enhance the EXs. Finally, Exs candidates were detected using histogram analysis [6] and remove falseEXs using multi-channel histogram analysis. To segment EXs from pre-processed image SBGFRL (Selective binary and Gaussianfiltering regularized level set) is applied [8]. Hariniet al.[10] used to add results of two operations for detecting EXs. The Firstoperation used top-hat transformation and illumination correction for background exclusion and other operation uses bloodvessels removal, segmentation and logical „&" operation. Convert retinal fundus image into a CMY color model and then extractmagenta component [11] to detect Exs. On magenta component applied thresholding base on a standard derivation of a magentacomponent to convert an image into binary image and then remove optic disc portion from a binary image. Finally to extract EXsapply morphological operation on the binary image. Morphological operation and segmentation techniques [12] are used to detectEXs.

4)MAs and HMs Detection: Microaneurysms and Hemorrhages are first DR symptoms. If we detect the earlier stage then we can reduce the blindness causes from DR. Two methods are used for detecting Mas and HMs [1]. The first method is Semi-automated DR detection using Eigenvalue of Hessian matrix Analysis. The second method is automatic DR detection using Eigenvalue of hessian matrix analysis, image processing techniques, and SVM classifier. DMPT-SC (Dynamic template scheme) is used to extract MAs [3]. They extract shape, pixels, algebra, and other features for template matching process. And also they used DCSAWS (Adaptive weighted scoring algorithm with distribution character based scoring scheme) to achieve high sensitivity [3].Canny edge detection is applied to a pre-processed image and then applied Hole Fill MATLAB technique [4].

Finally to extract MAs they subtract Hole Filled image and canny edge detected image. All features like blood vessels, optic disc and exudates aredetected using different image processing techniques [2, 6, 8 and 10]. And then detection of MAs take a place using differentmethods. Objects which are equal or smaller than 10 pixels detected as a MAs [2]. Matched filter and thresholding techniques were used to detect MAs and HMs [6]. INPAINTING technique is used to remove all the features from the pre-processed imageand then extended minima transform is used to detect MAs [8]. Pre-processing techniques were used to enhance the image andthen remove blood vessels from a preprocessed image, and then they created feature vector using gray level co-occurrencematrix, wavelet features and statistical features like area and shape. And used two classifier SVM and LMNN to detect MAs [5]. Methods like green channel, histogram equalization, and morphological operations are used to detect MAs and HMs in [7]. Bloodvessels are removed using average retina color and then disc based dilation is applied to highlight MAs [11]. Edge detection andHole Filling technique are applied to the dilated image. And subsequently, process these images to get possible MAs.Preprocessing techniques were brightness correction, gamma correction, green plane extraction, histogram equalization andthresholding is applied [13]. Subsequently, extract and remove blood vessels and Eigenvalue analysis to detect candidate region ofMAs.

C. Classification

Different tools and techniques are developed for the classification using some samples. The aim of this topic is to differentiate theplanes from where the testing samples are to be classified. The important aspect on which the accuracy of the system depends on theclassifier. Whatever the feature extraction is done to create a redundant and efficient data that has to be submitted to the classifierwhere the classifier is given some data for training purpose. Ten fundus images were finally given to the Support Vector Machine(S.V.M.) for training which consisted of positive and negative samples with specific labeling [1]. In addition, the principalcomponent analysis was also applied to the trained classifier.While similarly in [5] feature vector formed by the system is given to the S.V.M. They also used multi-layered feed forwardnetwork which is a Levenberg-Marquardt neural network algorithm to classify in MA's and Non-MA's. Similarly, S.V.M. classifieris used to create positive and negative planes and then test the data using trained S.V.M. classifier [10]. S.V.M. and decision tree were used but in two-phase DRNP detection and DRNP grade classification [11]. Two classifiers are used OneR classifier and BPNNs [12].

1) *SVM:* S.V.M. is a supervised training algorithm. Which used some training data set for the classification according to the number of class, and the dimension of the classifier can be increased. The classifier works in a way that the most relevant samples from the subplanes are used for creating the planes for the classification. SVM as a two-class classifier is also called linear classifier where only two dimensions or two classes are possible in which data samples are to be classified [21]. The SVM classifier takes the input of feature vector which is the result of the feature extraction techniques.

The result formed after the extraction of the feature are stored in feature set form that can be given to the SVM classifier for the training.

TheseSamples are also an aspect of the accuracy or the efficiency of the SVM classifier for the classification. SVM can also be used as the multiclass SVM where many "One-to-one SVM" classifier are combined to get a classifier for the required number ofthe class [21]. This is a way of increasing the dimension so that in our aspect of DR detection wherein [11] different grade ofthe disease are classified.

2) OneR: The OneR stands for one rule. This classification technique is simple yet accurate to classify. It creates set of rules for each predictor and picks the rule with a least total error [25]. For this purpose, it generates frequency table against the target. This generated rule is humanly understandable.

To generate one rule training data, attributes, and classes are taken as input andgenerate a set of rules for each attribute with the highest accuracy and minimum error.

3) *LMNNs:* Levenberg-Marquardt Neural Network is a multilayer feed-forward artificial neural network based on a mathematical model known as a Levenberg-Marquardt algorithm [22]. This mathematical model is used to update the weightsin the neural network. The algorithm is combined while training the algorithm shifts to the steepest descent algorithm [5].

4) *BPNNs:* Back Propagation Neural network is a multi-layered feed forward neural network. The training goes by updating the internal weight of the nodes for the precision. The training data set are taken for the training of the neural network. The training dataset has defined output which is given to the neural network. The difference between the actual and required output called an error is used to update the nodes in backpropagation in a way where the updating of the nodes travel from the output nodes to the internal nodes and the weights are the adjusted approximately and then subsequently the new output is achieved and the same process is repeated until the error is minimum or null [23].

5) Decision tree: This a very basic yet very useful in classification and regression hence also called a CART (classification andregression tree). In this, a tree is constructed in a way where a class attribute is taken as a leaf node and other attributes aretaken as an internal node [24]. The trees are constructed based on the probability of DR attributes denoting disease is taken intoaccount on basis of which the tree is constructed. This constructed tree can also be viewed in the form of statements like Pimplies Q. As decision tree is based on a probability of the attribute so a selection of the attribute is important and also in somecase number of an attribute is also important. On this both factor, the accuracy and the efficiency of this classifier aredependent. The different attribute from the training dataset's fundus image is to be extracted by feature extraction techniqueswhich have to be given for the construction of the decision tree and then similar attributes have to be extracted from the testingfundus image and given to the classifier were based on the constructed tree classification is done.

IV. RESULTS

For 874 subjects, the sensitivity of the device‟s rDR output to detect rDR was 96.8% (95% CI: 93.3%–98.8%) and specificity was 87.0% (95% CI: 84.2%–89.4%), with 6/874 false negatives, resulting in a negative predictive value of 99.0% (95% CI: 97.8%–99.6%), and positive predictive value of 67.4% (95% CI 61.5%–72.9%). Sensitivity for the device‟s rDR output to detect vtDR was 100% (95% CI: 96.1%–100%; i.e., no cases of vtDR were missed), and sensitivity for the device‟s rDR output todetect ME was also 100% (95% CI: 95.6%–100%; i.e., no casesof ME were missed by the rDR output). The AUC for thedevice‟s rDR index to detect rDR was 0.980 (95% CI: 0.968–0.992; Fig. 1; Table). The device‟s rDR output sensitivity to detect rDR, of 96.8%,was not statistically different from the previously publishedIDP sensitivity to detect rDR (P value 0.615), but its specificity,at 87.0%, was significantly better than that of IDP (P value <0.0001). The 6/874 false negatives are shown in Figure 2.For the device‟s vtDR output, sensitivity to detect vtDR was100.0% (95% CI: 96.1%–100.0%) and specificity was 90.8%(95% CI: 88.5%–92.7%), resulting in a negative predictive valueof 100.0% (95% CI 99.5%–100.0%), and positive predictivevalue of 56.4% (95% CI 48.4%–64.1%). The AUC for the device‟s vtDR index to detect vtDR was 0.989 (95% CI: 0.984–0.994;Fig. 1). We previously reported that the theoretical maximum AUCmeasurable on this specific dataset and reference standard hasa 95% CI of 0.939 to 0.972, which thus overlaps with the 95%CI of 0.968 to 0.992 of the measured AUC for the device‟s rDR index.1,18 Thirty-four subjects (4%) had at least one image that wasdeemed insufficient by the quality algorithm run outside thedevice..

V. CONCLUSION

[1] Diabetes is one of the fast-growing diseases in recent times. According to various surveys, a patient having diabetes has around 30% chances to get Diabetic Retinopathy (DR). DR has different stages from mild to severe and then PDR (Proliferative Diabetic Retinopathy). In the later stages of the diseases, it leads to floaters, blurred vision and finally can lead to blindness if it is not detected in the early stages. Manual diagnosis of these images requires highly trained experts and is time-consuming and difficult.
[2] Computer vision-based techniques for automatic detection of DR and its different stages have been proposed in the literature. We focused to classify all the stages of DR, especially the early stages, which is the major shortcoming of existing models. We proposed a CNN ensemblebased framework to detect and classify the Dr‟s different stages in color fundus images.
[3] We used the largest publicly available dataset of fundus images (Kaggle dataset) to train and evaluate our model. The results show that the proposed ensemble model performs better than other state-of-the-art methods and is also able to detect all the stages of DR.
[4] In future in order to further increase the accuracy of early-stage, we plan to train specific models for specific stages and then ensemble the outcome.

REFERENCE

[1] Journal article: IEEE Access-A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection.

[2] Journal article: IEEE Access-Automatic Analysis of Micro Aneurysms Turnover to Diagnose the Progression of Diabetic Retinopathy.

[3] SVM and Neural Network based Diagnosis of Diabetic Retinopathy -International Journal of Computer Applications (0975 – 8887) Volume 41– No.1, March 2012.

[4] A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection Received August 27, 2019, accepted October 4, 2019, date of publication October 15, 2019, date of current version October 29, 2019. Digital Object Identifier 10.1109/ACCESS.2019.2947484.AkaraSopharak a, Matthew N. Dailey b, BunyaritUyyanonvara a Sarah Barman c, Tom Williamson d, KhineThetNwe b and Yin Aye Moe b, "Machine learning approach to automatic exudate detection in retinal images from diabetic patients".*

[5] KanikaVerma, Prakash Deep and A. G. Ramakrishnan, Senior Member, IEEE, "Research Paper 4: Detection and Classification of Diabetic Retinopathy using Retinal Images."

[6] Wong Li Yun and Muthu Rama Krishnan Mookiah Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore 599489, "Detection of Diabetic Retinopathy Using K-Means Clustering and Self Organizing Map".

ISSN :25060

CREDIT CARD FRAUD DETECTION WITH DATA MINING AND MACHINE LEARNING APPROACH

[1] RUSHIKESH BASAVANT KOLTE, [2] KRISHNAKUMAR GAUTAM RATHOD, 3SUJEET MORE

[1,2] BE Student, Department of Information Technology TCOER , Pune
3Assistant Professor, Department of Information Technology, TCOER, Pune
E-mail: 1rushikeshkolte26@gmail.com

ABSTRACT

Nowaday's online payment gaining popularity because of easy and convenience use of ecommerce. It became very easy mode of payment. People choose online payment and e-shopping; because of time convenience, transport convenience, etc. As the result of huge amount of e-commerce use, there is a vast increment in credit card fraud also. Machine Learning has been successfully applied to finance databases to automate analysis of huge volumes of complex data. Machine Learninghas also played a salient role in the detection of credit card fraud in online transactions. Fraud detection in credit card is a big problem, it becomes challenging due to two major reasons–first, the profiles of normal and fraudulent behaviours change frequently and secondly due to reason that credit card fraud data sets are highly skewed. This paper research and checks the performance of Random Forest on highly skewed credit card fraud data. Dataset of credit card transactions is sourced from European cardholders containing 1 lakh transactions. These techniques are applied on the raw and pre-processed data. The performance of the techniques is evaluated based on accuracy, sensitivity, and specificity, precision.

Key words- Data Analysis, Fraud In Credit Card, Decision Tree, Random Forest, Machine Learning, Security.

I. INTRODUCTION

Credit card fraud is a growing concern with far reaching consequences in the government, corporate organizations, finance industry, In Today's world high dependency on internet technology has enjoyed increased credit card transactions but credit card fraud had also accelerated as online and offline transaction. As credit card transactions become a widespread mode of payment, focus has been given to recent computational methodologies to handle the credit card fraud problem. There are many fraud detection solutions and software which prevent frauds in businesses such as credit card, retail, e-commerce, insurance, and industries. Machine Learning is one notable and popular methods used in solving credit fraud detection problem. It is impossible to be sheer certain about the true intention and rightfulness behind an application or transaction. In reality, to seek out possible evidences of fraud from the available data using mathematical algorithms is the best effective option. Fraud detection in credit card is the truly the process of identifying those transactions that are fraudulent into two classes of legit class and fraud class transactions, several techniques are designed and implemented to solve to credit card fraud detection such as genetic algorithm, artificial neural network frequent item set mining, migrating birds optimization algorithm, comparative analysis of decision tree and random forest is carried out. Credit card fraud detection is a very popular but also a difficult problem to solve. Firstly, due to issue of having only a limited amount of data, credit card makes it challenging to match a pattern for dataset. Secondly, there can be many entries in dataset with truncations of fraudsters which also will fit a pattern of legitimate behavior. Also the problem has many constraints.

Firstly, data sets are not easily accessible for public and the results of researches are often hidden and censored, making the results inaccessible and due to this it is challenging to benchmarking for the models built. Datasets in previous researches with real data in the literature is nowhere mentioned. Secondly, the improvement of methods is more difficult by the fact that the security concern imposes a limitation to exchange of ideas and methods in fraud detection, and especially in credit card fraud detection. Lastly, the data sets are continuously evolving and changing making the profiles of normal and fraudulent behaviors always different that is the legit transaction in the past may be a fraud in present or vice versa. This paper evaluates two advanced machine learning, Decision tree and random forests and then a collative comparison is made to evaluate that which model performed best. Credit card transaction datasets are rarely available, highly imbalanced and skewed. Optimal feature (variables) selection for the models, suitable metric is most important part of mining to evaluate performance of techniques on skewed credit card fraud data. A number of challenges are associated with credit card detection, namely fraudulent behavior profile is dynamic, that is fraudulent transactions tend to look like legitimate ones, Credit card fraud detection performance is greatly affected by type of sampling approach used, selection of variables and detection technique used.

A. Motivation

The use of machine learning in fraud detection has been an interesting topic now days. A credit card frauddetection algorithm consists in identifying thosetransactions with a high probability of being fraud,based on historical fraud patterns. Machine learning, having three types, from that also the supervised andhybrid approach is more suitable for fraud detection.

B. Objectives

To formalization of the fraud-detection problem that realistically describes the operating conditions of frauds that everyday analyze massive streams of credit card transactions.

To design and assess a new technique that effectively addresses credit card frauds.

To Timely identification of fraudulent transactions can prevent the fraudsters from further committing such illicit crimes.

Problem Statement

To build credit card fraud detection system using machine learning algorithms. The major aim of this project is to perform a comprehensive review of different fraud detection methods and some innovative machine learning techniques.

II. LITERATURE REVIEW

In this section, we briefly review the related work on credit card fraud system and their different techniques.

In [1] this document proposes a new comparative measure of the comparison rules that reasonably represents the profits and losses due to fraud detection. A cost-sensitive method based on the minimum Bayes risk is presented using the proposed cost measure. Improvements of up to 23% are obtained by comparing this method and other latestgeneration algorithms. The data set for this document is based on the real-life transactional data of a large European company and personal data in the data is kept confidential. The accuracy of an algorithm is about 50%. The importance of this work was to find an algorithm and reduce the cost measurement. The result was 23% and the algorithm they found was the minimal risk of Bayes.

In [2] Several modern techniques based on sequence alignment, machine learning, artificial intelligence, genetic programming, data mining, etc. They have been developed and are still being developed to detect fraudulent credit card transactions. A solid and clear understanding of all these approaches is needed, which will undoubtedly lead to an efficient credit card fraud detection system. This document shows a survey of different techniques used in credit card fraud detection mechanisms and the evaluation of each methodology based on certain design criteria. An analysis of credit card fraud detection methods was performed. The survey in this document was based solely on detecting the efficiency and transparency of each method. The importance of this document was to conduct a survey to compare different credit card fraud detection algorithms to find the most appropriate algorithm to solve the problem.

In [3] A comparison was made between models based on artificial intelligence together with a general description of the fraud detection system developed in this document, such as the naive Bayesian classifier and the Bayesian network model, the clustering model. And finally, conclusions are drawn on the results of the model evaluation tests. The number of legal truncation was determined to be greater than or equal to 0.65, ie its accuracy was 65% using the Bayesian network. The importance of this document is to compare the models based on artificial intelligence together with a general description of the developed system and to establish the accuracy of each model together with the recommendation to create the best model.

In [4] Nutan and Suman on review on credit card fraud detection they have supported the theory of what is credit card fraud, types of fraud like telecommunication, bankruptcy fraud etc. and how to detect it, in addition to it they have explained numerous algorithms and methods on how to detect fraud using Glass Algorithm, Bayesian, networks, Hidden Markova model, Decision Tree and 4 more. They have explained in detail about each algorithm and how this algorithm works along with mathematical explanation. Types of machine learning along with classifications has been studied. Pros and cons of each method is listed.

III. EXISTING APPROACH

A lot of work has been done in this field thanks to its extensive use and applications. This section mentions some of the approaches that have been implemented to achieve the same purpose. These works are mainly differentiated from the technique for credit card fraud detection systems.

IV. PROPOSED APPROACH

In this system evidences from current as well as past ehaviour are combined. A fraud detection system is proposed that includes rule based filter, Dempster Shafer adder, transaction history database and Bayesian learner. In rule base the suspicion level of each incoming transaction is determined. Dumpster Shafer is used to combine multiple such evidences and an initial belief is computed. Based on this belief the transactions are classified as normal, abnormal or suspicious. The incoming transactions are initially handled by the rule base using probability values. After this the values are combined using Dumpster Shafer Adder. If the transaction is declared as fraudulent then it is handled by the card holder. If suppose the transaction is suspicious then it is fed in the suspicious table. The score of transaction is updated in the database with the help of machine learning classification. This architecture is flexible such that new kinds of fraud can be handled easily. With the help of learner the system can dynamically adapt to the changing needs.

System Diagram:

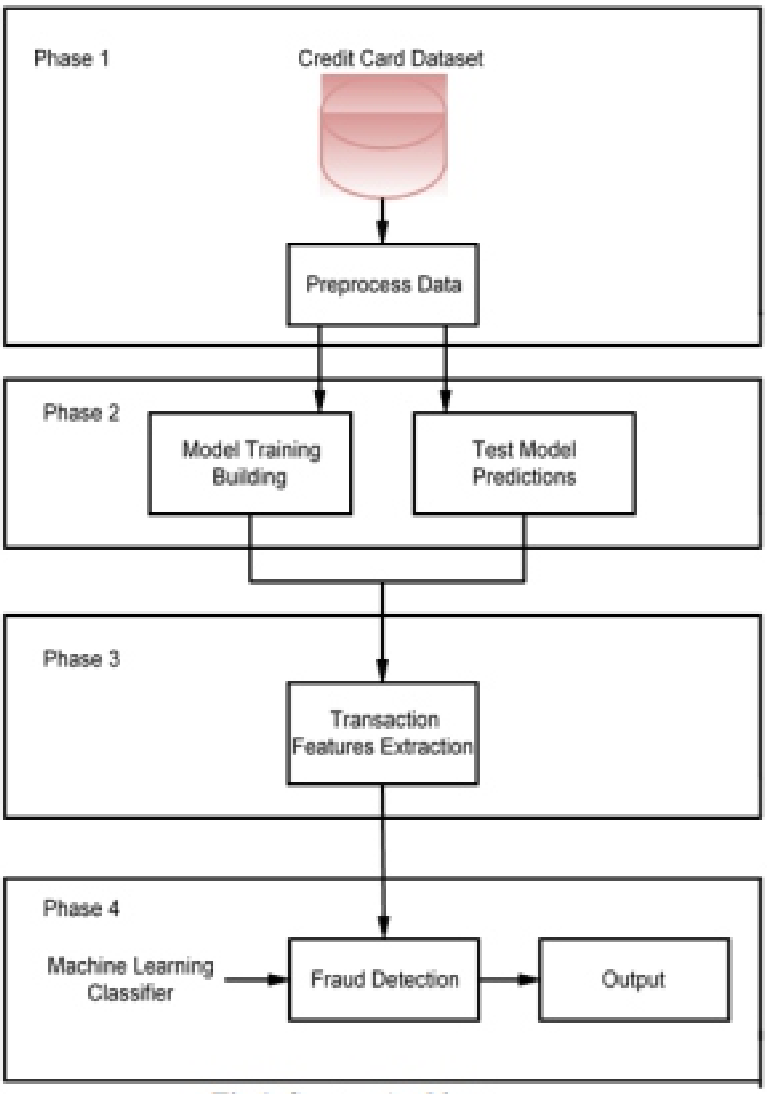


Fig 1. System Architecture

V. CONCLUSION

Credit card fraud detection is a fascinating domain. From this survey, we analyze that machine learning is the best compared to forecasting and classification. Machine learning techniques are mainly preferred in fraud detection, due to their high accuracy and detection rate. Even so, researchers find it difficult to achieve greater accuracy and detection speed. In addition, organizations are interested in finding ways to reduce costs and increase profits; you can find and select the method of previous studies.

REFERENCE

[1] Jain R., Gour B., Dubey S., A hybrid approach for credit card fraud detection using rough set and decision tree technique, International Journal of Computer Applications 139(10) (2016).

[2] Dermala N., Agrawal A.N., Credit card fraud detection using SVM and Reduction of false alarms, International Journal of Innovations in Engineering and Technology (IJIET) 7(2) (2016).

[3] Phua C., Lee V., Smith, Gayler K.R., A comprehensive survey of data mining-based fraud detection research. arXiv preprint arXiv:1009.6119 (2010).

[4] Bahnsen A.C., Stojanovic A., Aouada D., Ottersten B., Cost sensitive credit card fraud detection using Bayes minimum risk. 12th International Conference on Machine Learning and Applications

(ICMLA) (2013), 333-338..
[5] Carneiro E.M., Dias L.A.V., Da Cunha A.M., Mialaret L.F.S., Cluster analysis and artificial neural networks: A case study in credit card fraud detection, 12th International Conference on Information Technology-New Generations (2015), 122-126.
[6] Hafiz K.T., Aghili S., Zavarsky P., The use of predictive analytics technology to detect credit card fraud in Canada, 11th Iberian Conference on Information Systems and Technologies (CISTI) (2016), 1-6.
[7] Sonepat H.C.E., Bansal M., Survey Paper on Credit Card Fraud Detection, International Journal of Advanced Research in Computer Engineering & Technology 3(3) (2014).VarrePerantalu K., BhargavKiran, Credit card Fraud Detection using Predictive Modeling (2014).
[8] Stolfo S., Fan D.W., Lee W., Prodromidis A., Chan P., Credit card fraud detection using meta-learning: Issues and initial results, AAAI-97 Workshop on Fraud Detection and Risk Management (1997).
[9] Maes S., Tuyls K., Vanschoenwinkel B., Manderick, B., Credit card fraud detection using Bayesian and neural networks, International Journal of Pure and Applied Mathematics Special Issue 836Proceedings of the 1st international naiso congress on neuro fuzzy technologies (2002), 261-270.[10] Chan P.K., Stolfo S.J., Toward Scalable Learning with Non Uniform Class and Cost Distributions: A Case Study in Credit Card Fraud Detection, In KDD (1998), 164-168.

Instructions for Authors

Essentials for Publishing in this Journal

1 Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.

2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.

3 All our articles are refereed through a double-blind process.

4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

All articles for this journal must be submitted using our online submissions system. http://enrichedpub.com/ . Please use the Submit Your Article link in the Author Service area.

Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd**. The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page.

Type of Articles

Classification of articles is a duty of the editorial staff and is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).
2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);
3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);
4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);
2. Informative contribution (editorial, commentary, etc.);
3. Review (of a book, software, case study, scientific event, etc.)

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250-Word abstract should be placed between the title and the keywords with the body text to follow. Besides an abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

Keywords

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

Acknowledgements

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

Tables and Illustrations

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the text and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

Citation in the Text

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.
The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.

<u>**Address of the Editorial Office:**</u>

Enriched Publications Pvt. Ltd.
S-9,IInd FLOOR, MLU POCKET,
MANISH ABHINAV PLAZA-II, ABOVE FEDERAL BANK,
PLOT NO-5, SECTOR -5, DWARKA, NEW DELHI, INDIA-110075,
PHONE: - + (91)-(11)-45525005