Retina Screening Condition using CNN

SUREKHA A[1], ATCHYUTH KUMAR BOMMALA[2], SANDEEP JAMI[3], VASUDHA VAISHNAVI DULAM[4], CHANDINI PENTAKOTA[5]

[1] *Professor, Department of Information Technology, Anil Neerukonda Institute of Technology and Sciences, Sangivasala, Visakhapatnam, Andhra Pradesh, India*

[2, 3, 4, 5] *Department of Information Technology, Anil Neerukonda Institute of Technology and Sciences, Sangivasala,Visakhapatnam, Andhra Pradesh, India.*

Abstract— Retinopathy arises from poorly managed chronic diabetes and, if left untreated, can lead to complete vision loss. Thus, early detection and treatment are vital to mitigate its serious consequences. Manual diagnosis by eye specialists is time-intensive and burdensome for patients. An automated system offers a faster detection method, facilitating prompt treatment and reducing ocular complications. This research proposes a ML technique to extract features such as exudates, hemorrhages, and capillary aneurysms, utilizing a hybrid classifier comprising support vector machine, Logistic regression, random forest, and k-nearest neighbour multilayer perceptron network. The experiments resulted in an accuracy rate of 82%, with a precision score of 0.8120, a recall score of 0.8115 approximately, and an f-measure score of 0.8028 for the hybrid approach.

Index Terms— CNN classifier, Random Forest, Voting

I. INTRODUCTION

One type of retinopathy is grave complication of diabetes mellitus, posing a significant threat to global visual health due to potential to cause vision loss and blindness. With approximately one-third of the estimated two hundred and 50 million people globally are afflicted with diabetes exhibiting signs of retinopathy, the urgency of early detection cannot be overstated. Furthermore, projections indicating a substantial rise in affected individuals—from one hundred and twenty six million in 2010 to an anticipated one hundred and ninty million by 2030—underscore the urgent requirement for efficient procedures.Manifesting with symptoms such as blurred vision, darkened areas in vision, and eye floaters, diabetic retinopathy progresses through stages, with Non-Proliferative Diabetic Retinopathy (NPDR) representing a preliminary phase marked by the appearance of hemorrhages, capillary aneurysms, and exudates. While traditional physical examinations remain vital for diagnosis, their manual nature presents challenges in terms of time consumption and patient burden, emphasizing the imperative for automated detection systems. This paper addresses the critical need for efficient retina through the evolution of an automated computer-aided system.Leveraging ML techniques, particularly a hybrid model combining Support Machine (SVM),Convolational neural networks (CNN), the system that is suggested aims to extract key features indicative of retinopathy. By focusing on identifying hemorrhages, capill aryaneurysms, and they used to exudates in the retina, the system streamlines the detection process, offering a more efficient and effective approach to diagnosing diabetic retinopathy. This innovative utilization of computational algorithms promises to revolutionize diabetic retinopathy diagnosis, facilitating early intervention and ultimately improving patient outcomes.

II. LITERATURE REVIEW

Using the MESSIDOR dataset, [Farrikh Alzami, 2019] described a system for grading retina fractal analysis-based and random forest. Following the segmentation of the images by their framework, the fractal dimensions are computed as features. They had trouble telling the difference between moderate and severe retinopathy..

[Qomariah 2019] presented an self_operated system for classifying Support vector machines (SVM) and concurrent neural networks (CNN) are used to analyse normal retinal pictures and retinalopathy. Exudates, haemorrhage, and capillary aneurysms were among the characteristics. The suggested system was split into two halves by the writer: the initial section used SVM for classification, while the second section used neural network-based feature extraction .

Kumar (2018) proposed an improved a technique to identify diabetic retinopathy that uses color fundus pictures from the DIARETDB1 dataset to determine the number and area of capillary aneurysms.The fundus pictures and morphological procedures were all part of the pre-processing. Using an exponential SVM, the analysis of the principal components was conducted. contrast-limited adaptive histogram equalisation (CLAHE), morphological processes, and applied averaging filtering for regarding the designation and classification of capillary aneurysms.

In 2018, Mohamed Chetoui presented a method for identifying diabetic retinopathy. using different texture features and ML classification models. Two features, hemorrhage and exudates, were extracted using local ternary patterns (LTP) and local energy-based shape histograms (LESH). Feature vectors from LTP and LESH were utilized for instruction and classify the extracted histogram using SVM.

A technique for the fuzzy C-means-based feature extraction and SVM classification of retinopathy was proposed by [S Choudhury, 2016]. Using Arthematical morphology and a top- hat filter, blood vessel extraction was carried out. Exudates and etinal vascular density were chosen as the characteristics.Fuzzy C-means segmentation accustomed to achieve exudate extraction. To map SVM kernel space onto the training data, a The radial basis function of Gaussian employed.

[Sangwan, 2015] detailed a method that takes advantage of exudates, hemorrhage, and coronary arteries to differentiate between different phases of retinopathy. Using picture pre- processing, features were taken out and inputted into the neural network.The photos were classified into three groups using SVM-based training: mild, moderate non-proliferative diabetic retinopathy, and proliferative diabetic retinopathy. On the other hand, the approach would not work as intended if the regions of exudate in the fundus images were larger than an optical disc.
.

A technique for morphology-based exudate identification from fundus photos in color was described in [Morium Akter, 2014]. The model made use of histogram equalization, grayscale conversion, thresholding, dilation, erosion, logical AND operation, and alteration of watersheds. An output containing the ranges of exudates impacted in the diabetic retinopathy was created by the system.

A technique for categorising non-proliferative retinopathy using a soft-margin SVM was presented by Handayani (2013). The hard exudate in retina fundus images accustomed to categorised diabetic non-proliferative retinopathy severity level. Athematic morphology was being utilized the hard exudates in segments. Haemorrhage and capillary aneurysms, however, were not hallmarks within the system.

[Saravanan, 2013] suggested an self-operated system for detecting red lesions indicative of retina based on capillary aneurysms using a GMM classifier. Features derived through the application of mathematical morphology filter-based methods, and supervised learning. Severity levels of candidate capillary aneurysms were detected in four stages.

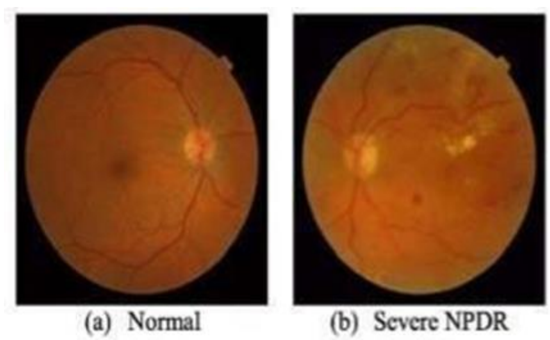
2011; Venkatalakshmi showed a hard exudate detection technique that is automated and uses color highlights and sharp edges as features. Precise edge identification, optic disc extraction, and color-based classification were among the techniques exploited throughout the detecting process. Training, as well as evaluation, occurred using the DRIVE and DIARETDB0 collections. The system's graphical user interface (GUI) was created using MATLAB 7.8.

III. METHODOLOGY

A.Dataset

For this purpose study, the Kaggle Dataset for Retinopathy Detection was made publically available. The dataset includes 1000 photos with and 1000 without retinopathy. retinopathy detection databases for retina that are accessible to the general public. We took 122 photos with retinopathy and 120 photos approximately without the condition from this pool. The selected aberrant pictures show clear characteristic such hemorrhages, exudates, and capillary aneurysms, which are important signs of retinopathys.

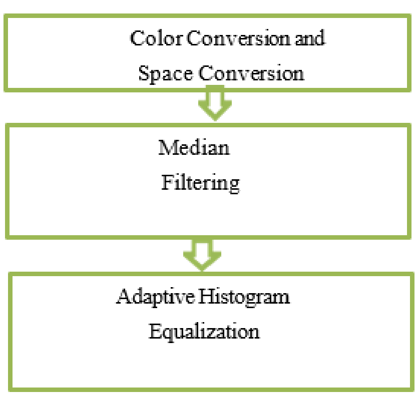
The determination of retina presence hinges on several factors, including the appearance, number, spread, and size of exudates, capillary aneurysms, and hemorrhages, as illustrated in Fig. 1. Exudates manifest as bright areas with a yellowish appearance, showcasing a color variance from the optic disc color. These formations often stem from ruptured blood vessels containing lipid. Meanwhile, hemorrhages result from ruptured capillary aneurysms in the red blood arteries. The spread of exudates and hemorrhages is particularly pronounced in vary retina images, indicative of the advanced phase of the condition.



a. Pre-processing

In the pre-processing stage, we initially converted from the dataset pictures to the HSV color space. This conversion facilitates the articulation of yellow-colored exudates from RGB images, as the RGB to HSV conversion enhances color differentiation. Subsequently, we performed edge zero padding, median filtering, and adaptive histogram equalization. Fig. 2 depicts images before and after pre-processing.

Steps:



The above figure describes about color conversion and the space conversion.This typically refers to transforming the raw data captured by the retinal imaging device in format format for suitable image manipulation as well as analysis.The image processing contains padding which refers to adding extra pixels around the borders of an image.Zero-Size padding, as the name suggests and involves in adding a zero-value pixels around the image.But in the median filter by a popular method for reducing images. Within the framework of retina images processing the images from the various resources like noise including speckle noise that causes by the image process or artifacts when they introduced during data acquisition.

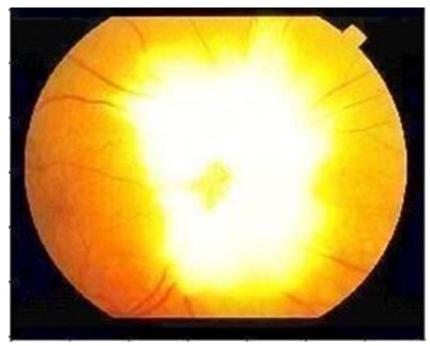
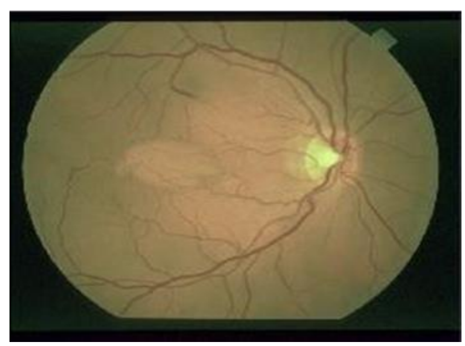


Fig 2 a) Image with exudates before pre- processing



b) Image without exudates before pre-processing

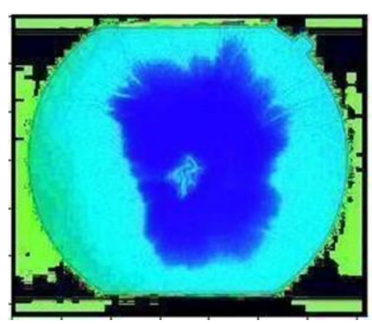
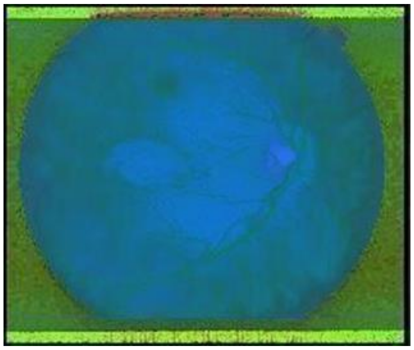


Fig 3 a) Picture including with exudates after pre-processing



b) a) Image with exudates after pre-processing

b. Image Segmentation

Following pre-processing, exudate segmentation entailed smoothing, masking, and bitwise AND operations. Smoothing was applied to eliminate frequency noise of high spatial ,while masking targeted yellow-colored exudates and the optic disc. Bitwise AND operations were employed for image manipulation, effectively isolating essential features. Similarly, hemorrhages and capillary aneurysms segmentation involved median blurring, thresholding, image erosion, and dilation. Fig 4 showcases images after segmentation processes.

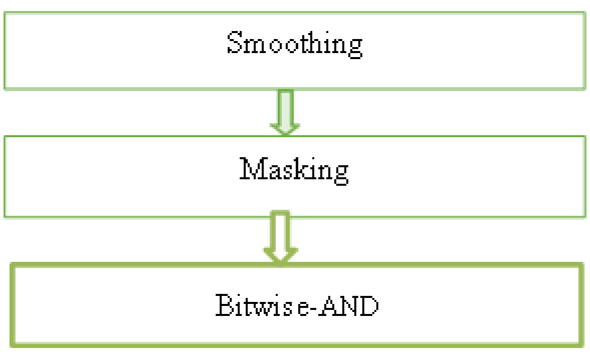


Fig: 4 Flowchart of exudate segmentation

Using the image processing technique of masking, we can alter a bigger image by defining a small 'image piece'. Here, we are using blue ([0,0,0255]) to mask the optic disc and yellow-colored exudates ([60,255,255]) in the smoothed image. When manipulating images, bitwise AND operations are employed to extract the image's main components.We can manipulate the larger Images by defining the smaller image pieces.

Bitwise techniques are useful for obscuring images. These operations used to facilitate image production. The qualities of the supplied photos may be improved by these operations. Here, we combine the image with the masked image to remove portions of the original picture other than the optical disc and exudates.

Figure 4: Following exudate segmentation, aberrant and the typical pictures are displayed.

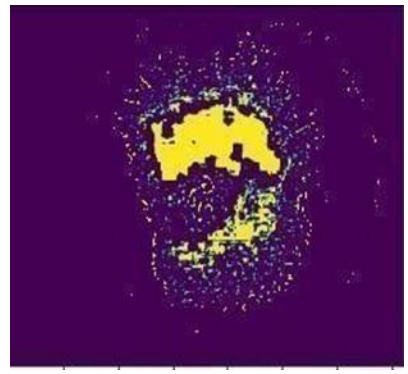


Fig4a) Images with Abnormal Segmented exudates

b) Regular pictures devoid of exudates

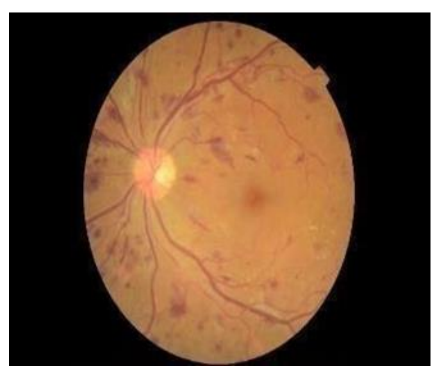


Fig 5a) Images with Abnormal micro aneurysms and haemorrhages

b) Images without Normal hemorrhages

Fig 5. c) Split-Open wounds and capillary aneurysmswith Abnormal images

d) Images with Normal after segmentation.

C.Classification

Random Forest and the k-nearest neighbours (KNN) algorithm were employed in categorization tasks. KNN is a supervised machine learning algorithms employed a straightforward and efficient method. relies on distance metrics to classify data points. Random Forest, on the other hand, employs a multitude of decision trees to collectively produce more accurate predictions. Additionally, a Voting mechanism was utilized to combine predictions from multiple classifiers, further enhancing classification accuracy.

Random Forest: The random forest is a technique for group learning that includes a multitude of separate decision trees. Each decision tree is structured with the top of its root and branches stemming downwards, with condition/internal nodes determining the splits. The terminal nodes, or leaves, represent the final decisions. The underlying principle of random forest lies in the concept of ensemble learning, where a substantial amount of relatively uncorrelated models (trees) collectively operate as a committee, often surpassing models with individual constituents.

Voting: Voting is a basic technique for forecasts derived from various machine learning techniques. Initially, several standalone models are created using the training dataset. A voting classifier is then employed to combine these models and average their predictions when someone presents new information for prediction. The predictions of the person sub-models can be weighted manually or heuristically to adjust their influence on the final prediction and it might be classified into two types hard voting and soft voting.In hard voting each model in the ensemble makes a prediction and the last prediction is determined by a majority vote.

Additionally, Convolational Neural Network(CNN) is a non- parametric classification algorithm that captures the concept of similarity, frequently called distance or proximity. A smaller distance implies a higher degree of similarity, guiding the classification process based on nearest neighbors.

IV. RESULTS AND DISCUSSION

The ability of a test is sensitivity to accurately identify individuals those are ill.TP / (FN+TP) equals sensitivity. The capacity of a test to correctly eliminate individuals who do not have a disease condition is known as specificity.TN /(FP+TN) equals specificity. A true positive (TP) occurs when a test yields a positive result and patient is able to identify the illness.When a test comes back negative and a individual is not given a diagnosis, this is often called the true negative (TN).When a test is positive but the subject is unable to communicate it, the condition is often a false positive (FP) False negative (FN) refers to a situation in which an individual may have a bad outcomes. SVM yields an accuracy of 68%.The precision of the KN classifier is 76%, whereas the haphazard woodland achieves 90%.Following the three classifiers votes, the testing set yields an accuracy of 82%. A 0.8619 accuracy score, a recall score f- measure score of 0.8028 and an f-value of 0.8116, 4 were acquired using the hybrid approach. That means that 36 of the 49 test samples yielded accurate predictions.

CONCLUSION

There are exudates, haemorrhages, and micro aneurysms this method that was recommended. The bitwise extraction, masking, and smoothing for exudate detection AND are carried out, leading to improved exudate extraction and computation.

Morphological procedures like opening the quantity of and how many exudates detect hemorrhages and micro aneurysms. Here, operators for erosion and dilation are carried out. To identify retina, count the quantity of MA events, the quantity of hemorrhages, and the how many number of secretions that happened in the image. This will allow us to ascertain the image's condition. Following that, characteristics are calculated. and sent into the Random Forest, KNN, and SVM classifiers. Three classifiers are voted upon to ascertain the final forecast. Therefore, it immediately determines if the disease grade is normal or abnormal according to the retrieved feature. Consequently, early identification and diagnosis people with Diabetes-related retinopathy can minimize the gravity of the condition and prevent blindness.

REFERENCES

[1] Farrikh Alzami, Abdussalam, Rama Arya Megantara and Ahmad Zainul Fanani, ―Diabetic Retinopathy Grade.

[2] Classification using Random Forest and Fractal AnalysisTM, which have been studied in relation to the retina pictures.The International Seminar on Application of Information Technology was initiated.

[3] ―Classification of the 12th International Conference on Information and Communication Technology and System Focused on retinopathy and Normal Retina Images utilising CNN and SVMTM in 2019.

[4] Detecting Diabetic Retinopathy by Determining the Area and Quantity by Shailesh Kumar and Basant Kumar's 5th International Conference on Signal Processing and Integrated Networks, 2018 of Micro-aneurysms from Color fundus pictures.

[5] Detecting Diabetic Retinopathy through Texture Analysis and Machine learning.Mohamed Chetoui, Moulay A. Akhloufi, and Mustapha Kardoucha, IEEE Canadian gathering on Electrical and Computer Engineering, 2018.

[6] The International Conference explores the application of Support Vector Machines in extracting features and classifying Diabetic Retinopathy ,specifically in conjuction with Fuzzy C means Communication and Signal Processing ,April 2016.S Bandyopadhyay,S Choudhury,SK Latin , DK Kole C.

[7] Vishal Sharma and Surbhi Sangwan at the International Conference on Computer and Computational Sciences, "Identification of different stages of diabetic retinopathy".

[8] Diabetic Retinopathy: Exudate Detection drawn from International Conference on Electrical Engineering and information and Communication Technology, 2014's Morphology from Colorful fundus pictures.

[9] Improving Reina Image Analysis to identify Diabetes Retinopathy: A Graphical User Interface (GUI) , Prof.

[10] B. Venkatalakshmi, V. Saravanan, and G. Jenny Niveditha, 2013.A. Narayanaswamy, S. Venugopalan, K. Widner, T. Adams, J. Cuadros, V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, Creating and verifying a deep learning technique to detect diabetic fundus opathy pictures .

[11] R.GargeyaandT.Leng,―Automated identification of retinopathy using deep learning,‖962–969, 2017.

[12] B. Graham, ―Kaggle retinopathy detection competition report.‖And they produced the link as shown below the https://kaggle2.blob.core.windows.net/forum and message of the attachments/88655/2795/competition report.pdf/,August6,2015 accessed May 20,2018.