DIABETIC RETINOPATHY DETECTION AND CLASSIFICATION USING MACHINE LEARNING

Dr. D. Lissy., MCA.,M.Phil.,Ph.D.,
Assistant Professor, Department of Computer Science and Engineering,
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology,
Chennai, India.

Dr. D. Suja Mary., M.Sc., M.Phil., Ph.D.
Assistant Professor, Department of Computer Applications,
St. Anne's Arts and Science College,Chennai-110, India

Dr. R.S. Dhivya
Assistant Professor, Department of Computer Science
Vallal P.T.Lee.Chengalvaraya Naicker Arts and Science College,
Chennai, India

Dr. H. Mohana., MCA., M.Phil., Ph.D.
Assistant Professor, Department of Computer Science,
St. Anne's Arts and Science College,Chennai-110, India

ABSTRACT

Diabetic retinopathy is an eye disease medical condition that caused by uncontrolled chronic diabetes and it can be causes complete blindness if not timely treated. The tiny sight vision problem occur on the retinal blood vessels. The early medical detection of diabetic retinopathy will cure is essential to prevent the severe side effects. Manual detection of diabetic retinopathy by ophthalmologist take plenty of time and patients need to suffer a lot at this time. The research work focuses to machine learning method for extracting three features like exudates, micro aneurysms and haemorrhages. The learning classification using hybrid classifier that the combination of various machine learning algorithms and multilayer perceptron network. The experiment result explores different king of learning algorithm styles and the highest accuracy values reduce the burden on ophthalmologists.

Keywords: Diabetic retinopathy, Machine Learning, Retinal fundus image, multilayer perceptron network, Image segmentation

I. INTRODUCTION

An automated method to detect the lesion belongs to diabetic retinopathy by applying digital image processing concepts in the field of medical diagnosis. The fundus image is taken by the fundus camera and it is given for pre-processing which reduces the noise in the image taken. After processing the image is undergone with feature extraction by the step-wise image processing morphological methods. In detecting abnormalities associated with fundus image, the images have to be Pre-Processed in order to correct the problems of uneven illumination problem, nonsufficient contrast between exudates and image background pixels and presence of noise in the input fundus image[1]. Pre-processing is an essential step in retinal image analysis which attenuates image variation by normalizing the original image with a reference model. Segmentation is generally the first stage in any attempt to analyze or interpret an image automatically. Segmentation bridges the gap between low-level image processing and high-level image processing[2]. Some kinds of segmentation technique will be found in any application involving the detection, recognition, and measurement of objects in images.

The proposed detection method to detect haemorrhages, exudates and micro aneurysms. For exudates detection green channel extraction, masking, smoothing, bitwise AND are done which results in better calculation and extraction of exudates[3]. For detection of haemorrhages and micro aneurysms, morphological operations are performed like opening. The presence of diabetic retinopathy is based on the appearance, number, spread and size, area of exudates, micro aneurysms, and haemorrhages. Then features are calculated and feed to SVM classifier[4]. This paper focuses on automated computer aided detection of diabetic retinopathy using machine learning hybrid model by extracting the features haemorrhage, micro aneurysms and exudates. The classifier used in this proposed model is the hybrid combination of SVM and KNN. The existing research experiments produced the highest accuracy values 82%. The proposed hybrid approach of this research work produced accuracy 98%, precision 97%, recall score 99% and f-measure score 99%.

II. BACKGROUND WORK

Diabetic Retinopathy (DR) is a major cause of irreversible visual impairment and blindness worldwide. This ethology of DR is due to chronic high blood glucose levels, which cause retinal capillary damage, and mainly affects the working-age population. DR begins at a mild level with no apparent visual symptoms but it can progress to severe and proliferated levels and progression of the disease can lead to blindness[5]. There are several standard DR grading systems such as the Early Treatment Diabetic Retinopathy Study (ETDRS)[6]. ETDRS separates fine detailed DR characteristics using multiple levels. The most critical state of DR is proliferative DR which causes new blood vessels to be formed and grown into the vitreous gel of the retina[7]. Machine learning classification algorithms namely Decision Tree, SVM and Naive Bayes are used in this experiment to detect diabetes at an early stage[8]. Experiments are performed on Pima Indians Diabetes Database (PIDD) which is sourced from UCI machine learning repository. The performances of all the three algorithms are evaluated on various measures like Precision, Accuracy, F-Measure, and Recall. Accuracy is measured over correctly and incorrectly classified instances[9]. Deep-learning (DL)-/machine-learning (ML)-based approaches make it possible to extract features from the images and to detect the presence of DR, grade its severity and segment associated lesions[10]. Convolutional Neural Networks (CNN) has been successfully applied in many adjacent subjects, and for diagnosis of diabetic retinopathy itself. However, the high cost of big labelled datasets, as well as inconsistency between different doctors, impedes the performance of these methods[11]. An automatic deep-learning-based method used for stage detection of diabetic retinopathy by single photography of the human fundus [12]. Convolutional Neural Networks (CNNs), a branch of deep learning, have an impressive record for applications in image analysis and interpretation, including medical imaging[13]. Network architectures designed to work with image data were routinely built already with useful applications and surpassed other approaches to challenging tasks like handwritten character recognition. However, it wasn't until several breakthroughs in neural networks such as the implementation of dropout, rectified linear units and the accompanying increase in computing power through Graphical Processor Units (GPUs) that they became viable for more complex image recognition problems [14]. On the Retinopathy Online Challenge's database, the method achieves a FROC score of 0.420 which ranks it fourth. On the Messidor database, when detecting images with diabetic retinopathy, that achieves an area under the ROC curve of 0.899, comparable to the score of human experts, and it outperforms state-of-the-art approaches[15].

III. DATASET AND METHODS

A. Dataset

This research paper used publicly available Kaggle Dataset for Diabetic Retinopathy Detection. The database was created with images taken from publicly available retinopathy detection datasets. The Kaggle dataset contain 1000 images with diabetic retinopathy and 1000 images without diabetic retinopathy[16]. From the total images we have chosen 122 images with diabetic retinopathy and 122 normal images. Chosen abnormal images contain exudates, haemorrhages, and micro aneurysms.

The is a pre-process dataset of the original available on Diabetic Retinopathy Detection. These images have been resized to 1024x1024 and they have been cropped in order to delete much of the black space. The Kaggle Diabetic Retinopathy (DR) dataset consists of 88,702 colour fundus images, including 35,126 samples for training and 53,576 samples for testing.

B. Pre-Processing

In image pre-processing, to find exudates, initially image from dataset is converted to HSV image. After image pre-processing, to segment exudates we have done smoothing, masking and bitwise AND. Smoothing is employed to remove high spatial frequency noise from image. This image analysis technique is a type of image segmentation that isolates objects by converting gray scale images into binary images. In the proposed method we are implementing hybrid classifier. That is we are using combination of three classifiers, Support vector machines, K nearest neighbours, Random forest.

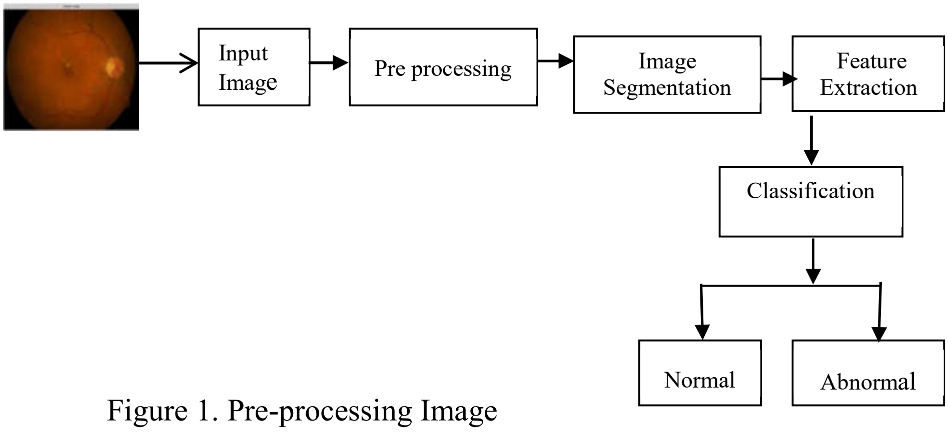

Figure 1. Pre-processing Image

Figure 1. shows image before pre-processing Colour space conversion is converting an image that is represented in one colour space to another colour space, the goal being to make the translated image look as similar as possible to the original. Red, Blue, Green channels in the given image to Hue, Saturation, Value. It is useful to extract yellow coloured exudates from RGB image when we convert RGB to HSV. Then edge zero padding, median filtering and adaptive histogram equalization is done.

C. Methodology

Image Segmentation

After image pre-processing, to segment exudates we have done smoothing, masking and bitwise AND. Smoothing is employed to remove high spatial frequency noise from image. Image blurring is achieved by convolving the image with a low-pass filter kernel. Masking is an image processing method in which we define a small 'image piece' and use it to modify a larger image. Here we are masking yellow coloured ([60,255,255]) exudates and optic disc in smoothed image with blue ([0, 0, 0255]) colour. Bitwise AND operations are used in image manipulation and used for extracting essential parts in the image. Bitwise operations help in image masking.

To segment haemorrhages and micro aneurysms median blurring, thresholding, image erosion and image dilation are performed. Image erosion and dilation are the morphological operations performed on image. Thresholding partitions an image into foreground and background. This image analysis technique is a type of image segmentation that isolates objects by converting gray scale images into binary images. Morphological Opening is defined as erosion followed by dilation. Opening can remove small bright spots and connect small dark cracks. This tends to open up gaps between features. Morphological erosion sets a pixel to the minimum over all pixels in the neighbourhood. Morphological dilation sets a pixel to the maximum over all pixels in the neighbourhood.

Feature Extraction

For binary classification, here we are using 2 features, i.e., number of exudates as first parameter and number of haemorrhages and micro aneurysms as second parameter. That is, we are counting number of white pixels from the segmented images and divide it by total number of pixels in the image.

Classification

In the proposed method we are implementing hybrid classifier. That is we are using combination of five classifiers, Support vector machines, K nearest neighbours, Random forest. Each classifier will classify the total 244 images into either normal or abnormal image. SVM classifier with kernel radial bias function and degree 3 is used. After obtaining the classifiers we have done voting as hybrid method. Training of dataset is done on five different classifiers and testing is done. Training and testing set are prepared in ratio 80:20. Support Vector Machine is a supervised machine learning algorithm can plot these each data item hat is n features as a point in n-dimensional space where the value of each feature represent the value of a particular coordinate in the n dimensional space. Then we classify the plotted data points into n classes by means of a hyper plane.

IV. EXPERIMENTAL RESULT AND DISCUSSION

Diabetes is a group of chronic diseases characterized by hyperglycaemia. Modern medical care uses a vast array of lifestyle and pharmaceutical interventions aimed at preventing and controlling hyperglycaemia. In addition to ensuring the adequate delivery of glucose to the tissues of the body, treatment of diabetes attempts to decrease the likelihood that the tissues of the body are harmed by hyperglycaemia. The importance of protecting the body from hyperglycaemia cannot be overstated; the direct and indirect effects on the human vascular tree are the major source of morbidity and mortality in both type I and type II diabetes. Diabetes complications are divided into two major categories:

A. Acute complications such as hypoglycaemia and comas resulting either form DKA or HHNS (as stated above) and

B. Chronic complications, either micro vascular (diabetic retinopathy, nephropathy, neuropathy) or macro vascular (coronary artery disease, peripheral arterial disease, and stroke)

Micro vascular is the most serious micro vascular diabetes complications are the eye complications. Diabetic patients are strongly advised to have an annual ophthalmic exam. Diabetic retinopathy is the leading cause of blindness in the working population of the Western world.

Diabetic retinopathy is generally classified as either background or proliferative. It is important to have a general understanding of the features of each to interpret eye examination reports and advise patients of disease progression and prognosis. Background retinopathy includes such features as small haemorrhages in the middle layers of the retina. They clinically appear as "dots" and therefore are frequently referred to as "dot haemorrhages." Proliferative retinopathy is characterized by the formation of new blood vessels on the surface of the retina and can lead to vitreous haemorrhage

Diabetic nephropathy is the leading cause of renal failure in the United States. About 20% to 30% of the patients with diabetes develop evidence of nephropathy. Initial treatment of diabetic nephropathy, as of other complications of diabetes, is prevention. Like other micro vascular complications of diabetes, there are strong associations between glucose control and the risk of developing diabetic nephropathy. Patients should be treated to the lowest safe glucose level that can be obtained to prevent or control diabetic nephropathy.

Diabetic neuropathy is recognized by the American Diabetes Association (ADA) as "the presence of symptoms and/or signs of peripheral nerve dysfunction in people with diabetes after the exclusion of other causes". As with other micro vascular complications, risk of developing diabetic neuropathy is proportional to both the magnitude and duration of hyperglycaemia, and some individuals may possess genetic attributes that affect their predisposition to developing such complications.

Macro vascular is the central pathological mechanism in macro vascular disease is the process of atherosclerosis, which leads to narrowing of arterial walls throughout the body. Atherosclerosis is thought to result from chronic inflammation and injury to the arterial wall in the peripheral or coronary vascular system Diabetes increases the risk that an individual will develop Cardiovascular Disease (CVD). Although the precise mechanisms through which diabetes increases the likelihood of atherosclerotic plaque formation are not completely defined, the association between the two is profound. Cardiovascular disease is a major complication and the leading cause of premature death among diabetic patients.

Original Image

In this proposed method haemorrhages, exudates and micro aneurysms are detected. For exudates detection green channel extraction, masking, smoothing, bitwise AND are done which results in better calculation and extraction of exudates. For detection of haemorrhages and micro aneurysms, morphological operations are performed like opening.

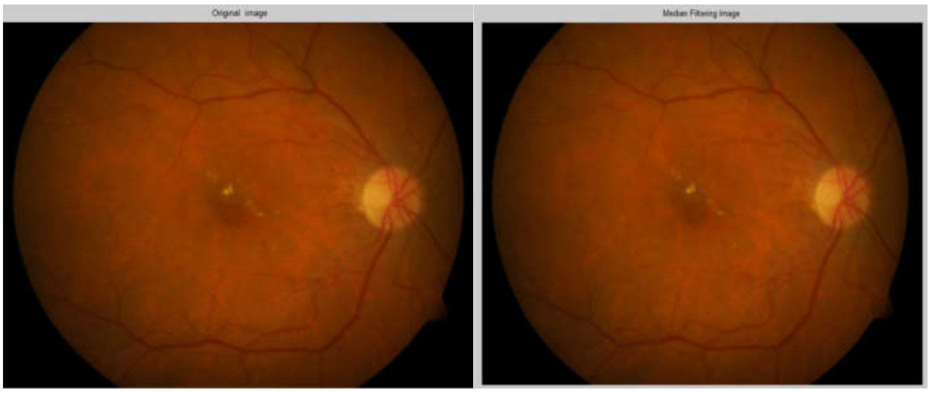

Figure 2 Original and Median Filtering Image

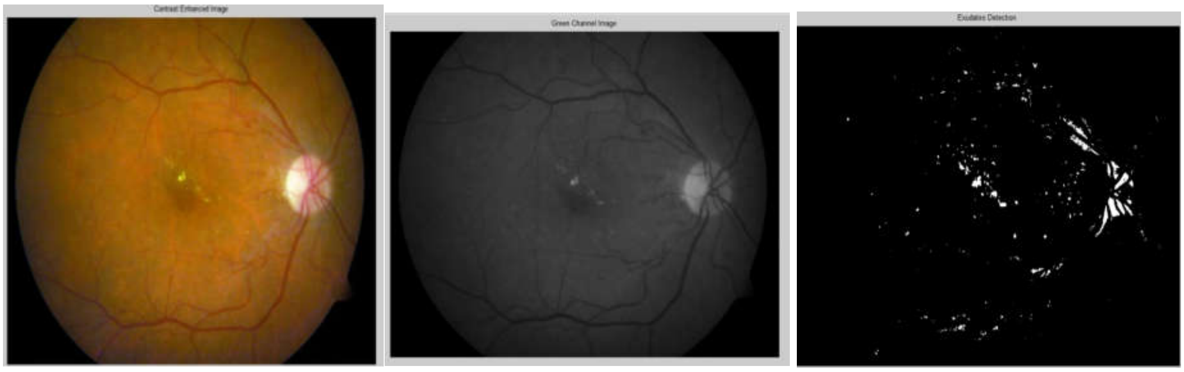

Figure 3. Contrast Enhanced Image and Green Channel Image and Exudates Detection

The presence of diabetic retinopathy is based on the appearance, number, spread and size, area of exudates, micro aneurysms, and haemorrhages is shown in figure2 and figure3.

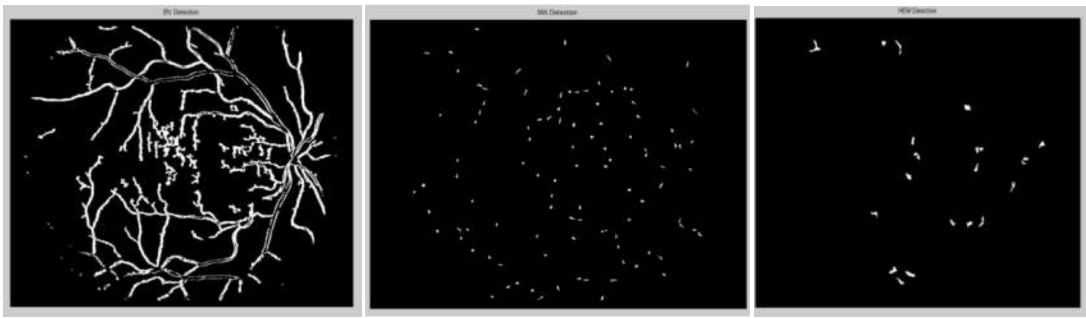

Figure 4. BV Detection and MA Detection and HEM Detection

The figure 4 shows that detection methods used in an image is to detect the Blood Vessels, Micro Aneurysms (MA) and Hemorrhages (HEM) Detection successfully. This methodology helps to find the disease Diabetic Retinopathy.

Many tests were conducted to select the best SVM kernels and optimize the hyper-parameters. SVM with RBF was the best performing followed by the polynomial kernel for the LTP accuracy. Table I gives the ACC (accuracy) and AUC (Area under the ROC curve) results for the different tests. We can see that LTP gives a very interesting performance. However, LESH is the best performing and obtain an ACC of 0.904 using SVM with RBF kernels. The same higher performance is obtained by LESH for AUC with RBF kernels which give 0.931. These results outperform LTP counterparts and show the robustness of LESH features in classifying DR. In the other hand, we can see that the LBP performance is lower than LTP and LESH.

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

$$F1 = \frac{2PR}{P + R}$$

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP + TN}}{\boldsymbol{TP + TN + FP + FN}}$$

Method	**Accuracy/%**	**Precision/%**	**Recall/%**	**F1/%**
Existing	88.73	95.06	92.93	92.66
Proposed	98.18	97.45	99.81	99.79

Table 1. Comparison table

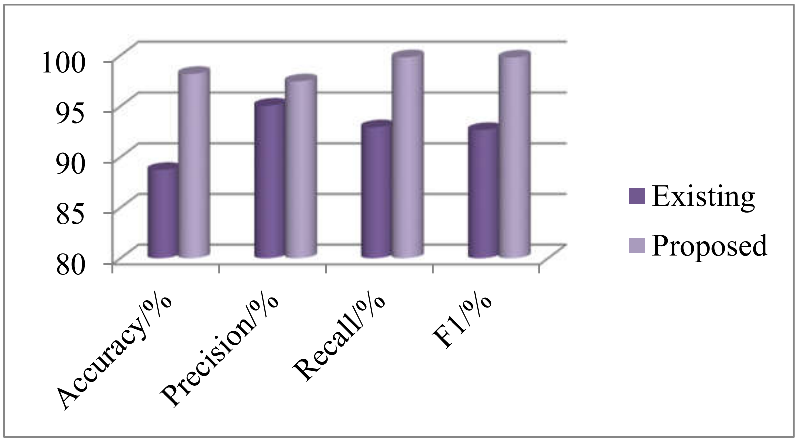

Figure 5. Comparision table chart

The figure 5 described the experimental result that given in the table 1. The proposed system results are efficient than the existing results and the accuracy score is increased.

V. CONCLUSION AND FUTURE WORK

In this proposed method haemorrhages, exudates and micro aneurysms are detected. For exudates detection green channel extraction, masking, smoothing, bitwise AND are done which results in better calculation and extraction of exudates. For detection of haemorrhages and micro aneurysms, morphological operations are performed like opening. Dilation and erosion operators are performed here. For diabetic retinopathy detection, count the number for MA occurred, count the number of haemorrhages occurred and count the number of exudates occurred in the image so we can decide the condition of image. Then features are calculated and feed to both SVM, KNN, Random Forest classifier. Voting of three classifiers is chosen as final prediction. So from the extracted feature it directly concludes the disease grade as normal or abnormal. So earlier detection and diagnosis of diabetic retinopathy help the patients from blindness and also the severe effects of disease can be decreases.

In the future, the proposed trained model will be used to build a mobile app with a user-friendly interface. Also this proposed framework will be applied to other medical contexts to compare their generality and versatility to predict the disease classes. The future works consider the train data in system, with more number of pre-processed data for getting higher accuracy result retinopathy. To train in network architecture and network parameters have been developed manually taking time-consuming and error-prone process. So the future workuse hyper-parameters optimization in machine learning and neural architecture in deep learning.

REFERENCES

[1] Shyamala Subramanian, Sashikala Mishra, "Machine Learning Styles for Diabetic Retinopathy Detection: A Review and Bibliometric Analysis", Big Data Cogn. Comput. 2022, 6, 154. https://doi.org/10.3390/bdcc6040154

[2] Revathy R1 , Nithya B S," Diabetic Retinopathy Detection using Machine Learning", (IJERT) Vol. 9 Issue 06, June-2020, ISSN: 2278-0181.

[3] Kaur J. Mitta D. A generalized method for the detection of vascular structure in pathologicalretinal images. Biocybernetics and Biomedical Engineering, Volume 37, Issue 1, 2017, Pages 184-200

[4] SohiniRoychowdhury, Dara D Koozekanani, Keshab K Parhi , "DREAM: Diabetic retinopathy analysis using machine learning", IEEE J Biomed Health Inform. 2014 Sep;18(5)pg:1717-28. doi: 10.1109/JBHI.2013.2294635.

[5] Ruchir Srivastava, Damon W. K. Wong, LixinDuan, Jiang Liu, "Red lesion detection in retinal fundus images using frangi-basedfilters", 37th Annual Internationational conference, IEEE 2015

[6] Vasudevan Lakshminarayanan, Hoda Kheradfallah, "Automated Detection and Diagnosis of Diabetic Retinopathy: A Comprehensive Survey", *J.Imaging* **2021**, *7*(9), doi:**https://doi.org/10.3390/jimaging7090165**

[7] M. Mohsin Butt, Ghazanfar Latif," Multi-channel Convolutions Neural Network Based Diabetic Retinopathy Detection from Fundus Images", 16th International Learning & Technology Conference 2019, Procedia Computer Science 163 (2019) 283–291.

[8] D. Sisodia, ''Prediction of diabetes using classification algorithms,'' Procedia Compute. Sci., Jan. 2018, vol.132, pg. 1578–1585,

[9] Farrikh Alzami, Abdussalam, Rama Arya Megantara and Ahmad Zainul Fanani, "Diabetic Retinopathy Grade Classification based on Fractal Analysis and Random Forest", International Seminar on Application for Technology of Information and Communication, 2019.

[10] Borys Tymchenko, Philip Marchenko, Dmitry Spodarets, "Deep Learning Approach to Diabetic Retinopathy Detection" International Conference on Pattern Recognition Applications and Method, ICPRAM 2020 , ISSN: 2184-4313, pg. 501-509,DOI: 10.5220/0008970805010509

[11] Adrian Galdran, José Dolz, Hadi Chakor, Hervé Lombaert, Ismail Ben Ayed, "Cost-Sensitive Regularization For Diabetic Retinopathy Grading From Eye Fundus Images ", MICCAI 2020 - OCTOBER 2, 2020

[12] Harry Pratta, Frans Coenenb, "Convolutional Neural Networks for Diabetic Retinopathy", International Conference On Medical Imaging Understanding and Analysis 2016, MIUA 2016, 6-8 July 2016, Loughborough, UK. DOI: 10.1016/j.procs.2016.07.014

[13] Nagaraja Gundluru,1Dharmendra SinghRajput,2Kuruva Lakshmanna," Enhancement of Detection of Diabetic Retinopathy Using Harris Hawks Optimization with Deep Learning Model", Hindawi Computational Intelligence and Neuroscience Volume 2022, Article ID 8512469, doi:https://doi.org/10.1155/2022/8512469

[14] Lama Seoud; Thomas Hurtut," Red Lesion Detection Using Dynamic Shape Features for Diabetic Retinopathy Screening", IEEE Transactions on Medical Imaging, Volume: 35, Issue: 4, April 2016, DOI: 10.1109/TMI.2015.2509785

[15] Deepti Sisodiaa , Dilip Singh Sisodia," Prediction of Diabetes using Classification Algorithms", Procedia Computer Science 132 (2018) 1578–1585.

[16] B. Graham, "Kaggle diabetic retinopathy detection competition report," https://kaggle2.blob.core.windows.net