Thesis/Project No. – CSER-20-30

Diabetic Retinopathy Classification from Retinal Images using Machine Learning Approaches

By

Tareq Mahmud

Roll No.:1507089

&

Indronil Bhattacharjee

Roll No.:1507105

Department of Computer Science and Engineering,
Khulna University of Engineering & Technology,
Khulna-9203, Bangladesh.

February 2020

Diabetic Retinopathy Classification from Retinal Images using Machine Learning Approaches

Tareq Mahmud

Roll No.:1507089

&

Indronil Bhattacharjee

Roll No.:1507105

A thesis submitted in partial fulfillment of the requirements for the degree of "Bachelor of Science in Computer Science & Engineering"

Supervisor:
Mr. Al-Mahmud
Assistant Professor
Department of Computer Science and Engineering,
Khulna University of Engineering & Technology.

_____
Signature

Department of Computer Science and Engineering,
Khulna University of Engineering & Technology,
Khulna-9203, Bangladesh.
February 2020

Acknowledgement

All praises go to Almighty Allah for showering His endless blessings and kindness on us. Without His desire, we could not have been in the position we are today. We would like to express our deep gratitude to Mr. Al-Mahmud, Assistant Professor, our thesis supervisor, for his patient guidance, enthusiastic encouragement and useful critiques of this thesis. He encouraged us with his useful and constructive recommendations throughout the planning of this thesis. We offer our sincerest regards to him for supporting us with his knowledge and experience whilst allowing us the room to think in our own way. His willingness to give his time so generously is very much appreciable. His blessing help and guidance, time by time, will carry us a long way in the journey of life on which we are about to embark.

Any constructive comments, suggestions, criticism from teachers as well as others will be highly appreciated and gratefully acknowledged.

-Authors

Abstract

Diabetic Retinopathy is one of the common eye diseases and is a diabetes complication that affects eyes. Diabetic retinopathy may cause no symptoms or only mild vision problems. Eventually, it can cause blindness. So early detection of symptoms could help to avoid blindness. In this thesis, we present some experiments on some features of Diabetic Retinopathy like properties of exudates, properties of blood vessels and properties of microaneurysm. Using the features, we can classify healthy, mild non-proliferative, moderate non-proliferative, severe non-proliferative and proliferative stage of DR. Support Vector Machine, Random Forest and Naive Bayes classifiers are used to classify the stages. Finally, Random Forest is found to be the best for higher accuracy, sensitivity and specificity of 76.5%, 77.2% and 93.3% respectively.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

According to WHO (World Health Organization) more than 347 million people are suffering from diabetes and it will be seventh prominent reason of death worldwide in 2030. Over the years, patients with diabetes tend to show abnormality in retina, due to emerging obstacle called Diabetic Retinopathy. People above 30 years having diabetes for more than 15 years, carry 78% chance of developing Diabetic Retinopathy. Diabetic Retinopathy is due of long-term standing of diabetic mellitus. Retinopathy means - damage of retina and as a result, the blood vessels become choked, leaky and grow arbitrarily. Diabetic Retinopathy is asymptomatic; it does not affect with view until it reaches at progress stage. Therefore, screening of Diabetic Retinopathy is crucial for type1 (Non-Proliferative) and type2 (Proliferative) diabetic patients as both types are at risk of Diabetic retinopathy.

1.1 Diabetic Retinopathy

People with diabetes can have an eye disease called diabetic retinopathy. This is when high blood sugar levels cause damage to blood vessels in the retina. These blood vessels can swell and leak. Alternatively, they can close, stopping blood from passing through. Sometimes abnormal new blood vessels grow on the retina. All of these changes can steal one's vision.

1.2 Stages of Diabetic Retinopathy

There are two main stages of diabetic eye disease.

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

1.2.1 Non-Proliferative Diabetic Retinopathy

Non-Proliferative Diabetic Retinopathy is the early stage of diabetic eye disease. Many people with diabetes have it. With NPDR, tiny blood vessels leak, making the retina swell. When the macula swells, it is called macular edema. This is the most common reason why people with diabetes lose their vision.

Also with NPDR, blood vessels in the retina can close off. This is called macular ischemia. When that happens, blood cannot reach the macula. Sometimes tiny particles called exudates can form in the retina. These can affect your vision too. If anybody has NPDR, his vision will be blurry.

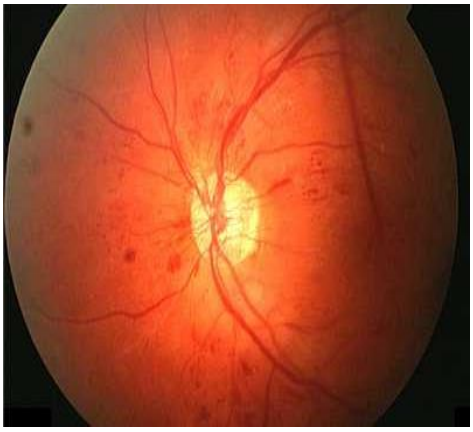


Fig. 1.2.1.1. A Non-Proliferative Diabetic Retinopathy affected eye

1.2.2 Proliferative Diabetic Retinopathy

PDR is the more advanced stage of diabetic eye disease. It happens when the retina starts growing new blood vessels. This is called neovascularization. These fragile new vessels often bleed into the vitreous. If they only bleed a little, you might see a few dark floaters. If they bleed a lot, it might block all vision.

These new blood vessels can form scar tissue. Scar tissue can cause problems with the macula or lead to a detached retina. PDR is very serious, and can steal both your central and peripheral (side) vision.

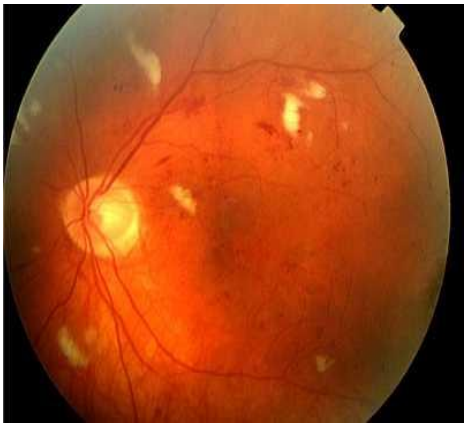


Fig. 1.2.2.1. A Proliferative Diabetic Retinopathy affected eye

1.3 Thesis Objectives

Our thesis objective is to detect diabetic retinopathy and classify whether it is a PDR or NPDR. Specific objectives of the thesis are:

- Process color fundus retinal images for Diabetic Retinopathy detection.
- Extract key features from the pre-processed images.
- Detect the presence of Diabetic Retinopathy.
- Classify whether the Diabetic Retinopathy is Proliferative or Non-proliferative.

1.4 Thesis Organization

Chapter 1 already discussed about Diabetic Retinopathy, existing methods related to Diabetic Retinopathy detection and thesis objectives.

Chapter 2 will discuss about survey of existing methods related to Diabetic Retinopathy detection and analyzes them.

Chapter 3 will discuss about the method for Diabetic Retinopathy detection from retinal fundus images.

Chapter 4 will discuss about experimental studies and results of the thesis.

Chapter 5 will discuss about concluding remarks of the thesis.

Chapter 2

Literature Review

Diabetic patients need regular screening because early detection of exudates could help to prevent blindness. However, manual examination by ophthalmologists takes time and the number of experts is not sufficient to meet the demand for screening. Given the limitations of manual screening, the prospect of automatic detection of retinal exudates, towards diagnosis and tracking the progress of a patient's treatment program, is enticing. There have been several attempts to solve this problem. Quite a few are based on thresholding and region growing:

2.1 Region Growing Based Segmentation Based Methods

- Liu et al. [4] detect exudates using thresholding and region growing. Their fundus photographs were taken with a non-mydriatic fundus camera then scanned by a flatbed scanner.
- Sinthanayothin et al. [6] report the result of an automated detection of diabetic retinopathy on digital fundus images using recursive region growing segmentation (RRGS). They measure performance on 10x10 pixel patches.
- Usher et al. [7] detect candidate exudate regions using a combination of RRGS and adaptive intensity thresholding. The candidate regions thus extracted are classified as exudate or non-exudate by a neural network.
- Kavitha and Shenbaga [8] propose median filtering and morphological operations for blood vessel detection. They use multilevel thresholding to extract bright regions assumed to be the optic disc or exudates. They detect the optic disc as the converging point of the blood vessels, then classify the other bright regions as exudates. The method performed poorly on low-contrast images. Thresholding and region growing methods are straightforward, but selecting threshold values, region seed points, and stopping criteria are difficult.

2.2 Clustering Based Methods

Clustering has also been proposed as a possible solution to the exudate detection problem:

- Osareh et al. [9] use fuzzy c-means clustering to segment colour retinal images into homogeneous regions, then train neural networks and support vector machines (SVMs) to separate exudate and non-exudate areas.
- Zhang et al. [10] use local contrast enhancement and fuzzy c-means clustering in the LUV colour space to segment candidate bright lesion areas. They use hierarchical SVMs to classify bright non-lesion areas, exudates, and cotton wool spots.

2.3 Image Processing Based Methods

The main difficulty with clustering methods is determining the number of clusters to use. A few other attempts are based on specialized features and morphological reconstruction techniques:

- Katarzyna et al. [11] detect candidate exudate regions with a watershed transformation and a marker and extract the optic disc based on geodesic reconstruction by dilation.
- Sanchez et al. [12] combine colour and sharp edge features to detect exudate. First they find yellowish objects, then they find sharp edges using various rotated versions of Kirsch masks on the green component of the original image. Yellowish objects with sharp edges are classified as exudates.
- Walter et al. [13] use morphological reconstruction techniques to detect contours typical of exudates.
- Acharya et al. [18] automatically identified normal, mild DR, moderate DR, severe DR, and PDR stages using the bispectral invariant features of higher-order spectra techniques and a SVM classifier.

2.4 Neural Network Based Methods

- Wang et al. [14] extract colour features then use a feed forward neural network to identify retinal lesions.
- Ege et al. [20] use a median filter to remove noise, segment bright lesions and dark lesions by thresholding, perform region growing, then identify exudates regions with Bayesian, Mahalanobis, and nearest neighbour (NN) classifiers. The system failed to detect exudates in low quality images.

2.5 Scope of this Research

Most of the related works classifies the disease with two classifiers, i.e. whether he/she has a healthy eye or has a diabetic retinopathy defective eye. The scope in our research is to split this classes into proliferative and non-proliferative stages of diabetic retinopathy. An eye can be classified into 5 classes as Healthy, Mild non-proliferative, Moderate non-proliferative, Severe non-proliferative and proliferative eye. Besides, extracting hand crafted features from raw images after different processing is one of the best scopes we have worked in this research.

In this research, image processing based method has been used. The objective of using this method is to get some features which are specifically responsible for the diabetic retinopathy. To get this features, several morphological functions and reconstructions has been used. Exudates, Blood vessels and Microaneurysms are taken into account, because these features are the most disastrous for eye diseases. From the study on the related works, it has been recognized that applying machine learning based classifiers for hand crafted features from raw images gives more accurate predictions. Based on the study, Random Forest, Support Vector Machine and Naïve Bayes are selected for the research.

Chapter 3

Methodology

Color fundus image of eye is provided as an input and by processing this, we get the result whether the eye is affected with Diabetic Retinopathy and stages of DR. To detect Diabetic Retinopathy, we preprocess the image and apply classification techniques. We divide the total process into 3 steps- preprocessing, feature extraction and Diabetic Retinopathy classification. Preprocessing includes green channel extraction, Contrast Limited Adaptive Histogram Equalization, dilation, morphological process, median filtering, thresholding etc. In the feature extraction phase, we extract several features like area of exudates, area of blood vessels, area of microaneurysm etc. Finally, in the classification phase, we will detect whether Diabetic Retinopathy is present or not. Moreover, if present, whether it is Mild, Moderate, Severe and PDR.

3.1 Diabetic Retinopathy Classification System

• **Input:**

– Colour fundus retinal images

• **Output:**

– Diabetic Retinopathy is not present, Mild, Moderate, Severe or PDR.

• **Process:**

Step 1: Take the input image

Step 2: Preprocess the input image

Step 3: Exudates detection

Step 4: Blood Vessels detection

Step 5: Microaneurysm detection

Step 6: Feature extraction

Step 7: Apply Random Forest Classifier

Step 8: Classify the Diabetic Retinopathy stages

Step 9: Detect whether it is Healthy, Mild, Moderate, Severe or PDR eye

3.2 Modular Flow of the Methodology

The modular flow of the proposed system is like the following-

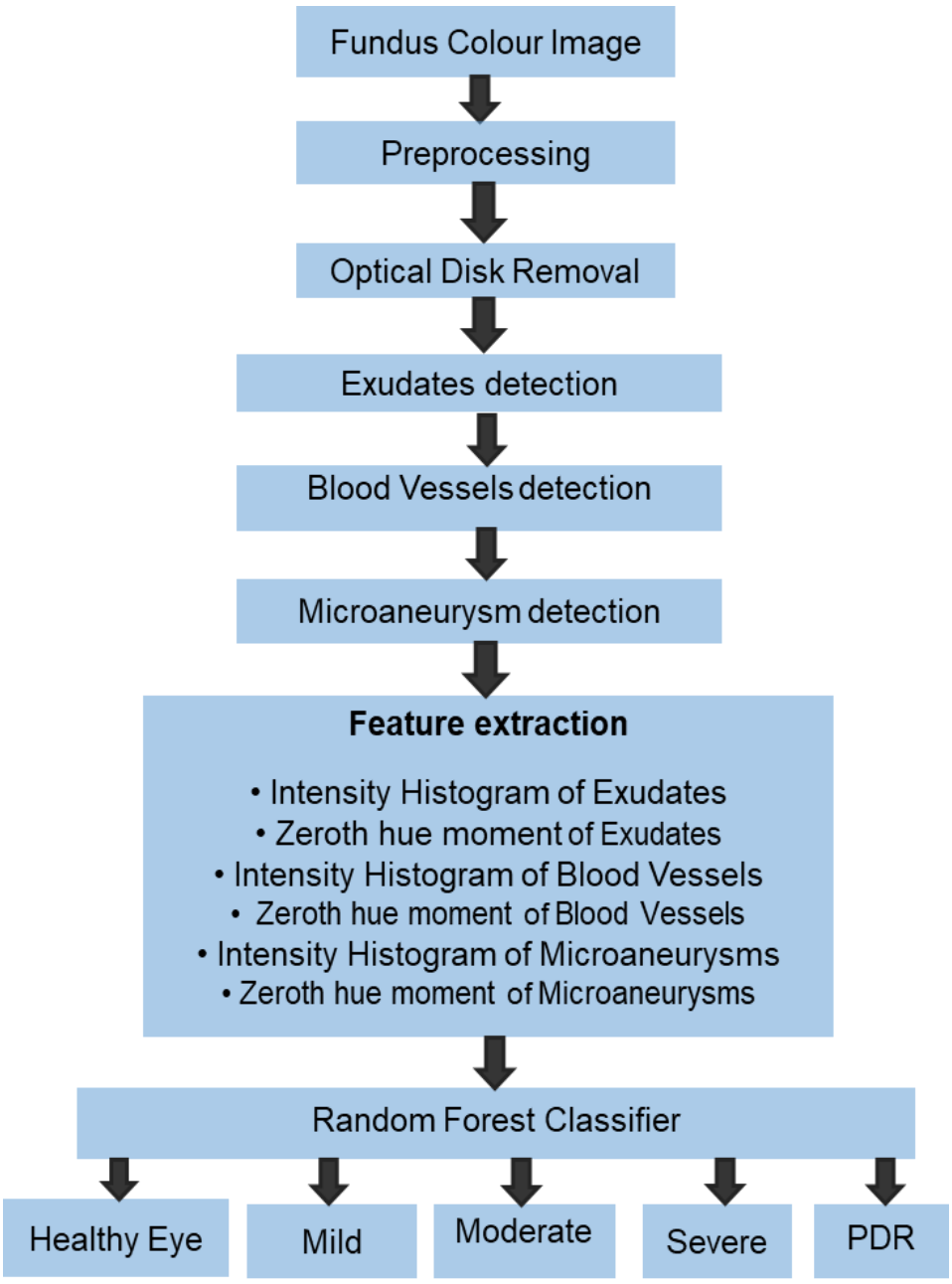


Fig. 3.2.1. The flow chart of the Diabetic Retinopathy Classification Process

3.3 Steps of Diabetic Retinopathy Detection

The proposed Diabetic Retinopathy Detection approach is described in the following three steps.

1. Preprocessing
2. Feature Extraction
3. Classification

3.3.1 Preprocessing

The Preprocessing has been done in two steps. First one is the general preprocessing, which is applied for all the images of the dataset. Then second one is the specific preprocessing according to the features to be extracted.

3.3.1.1 General Preprocessing

- **Resizing**

In this work, the sizes of the actual images are 1024x1024 pixels. As the dataset is huge in size, the images are stored in .jpeg format with size 300x300 pixels for reducing the computational time.

- **Green Channel Extraction:**

Pre-processing of fundus image is performed in order to improve the contrast. In order to enhance the contrast of the retinal images, some information is commonly discarded before processing such as the red and blue components of the image. Green channel is extensively used in pre-processing as it displays the best vessels/background contrast and greatest contrast between the optic disc and retinal tissue [2]. Red channel is relatively bright and vascular structure of the choroid is visible. The retinal vessels are also visible but show less contrast than green channel. Blue channel is noisy and contains little information.

- **Contrast Limited Adaptive Histogram Equalization:**

Contrast Limited Adaptive histogram equalization (CLAHE) is used for contrast enhancement. CLAHE computes several histograms of image and uses them to reallocate intensity value of image. Hence, CLAHE is more appropriate to improve regional contrast and edge enhancement in each region of image.

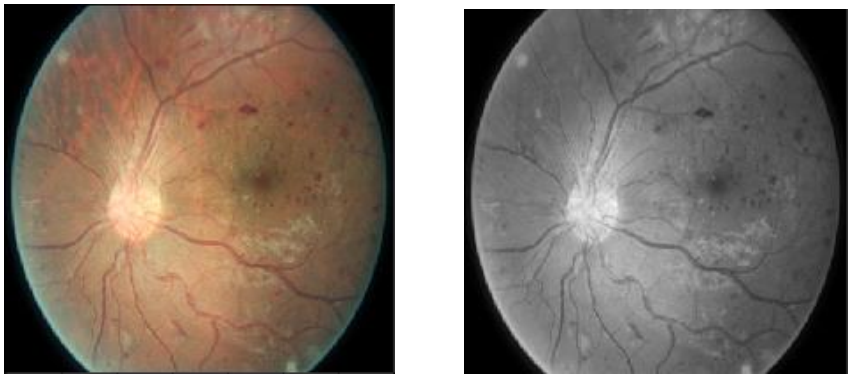
3.3.1.2 Specific Preprocessing

The specific preprocessing is performed for exudate detection by using morphological process, median filtering and thresholding.

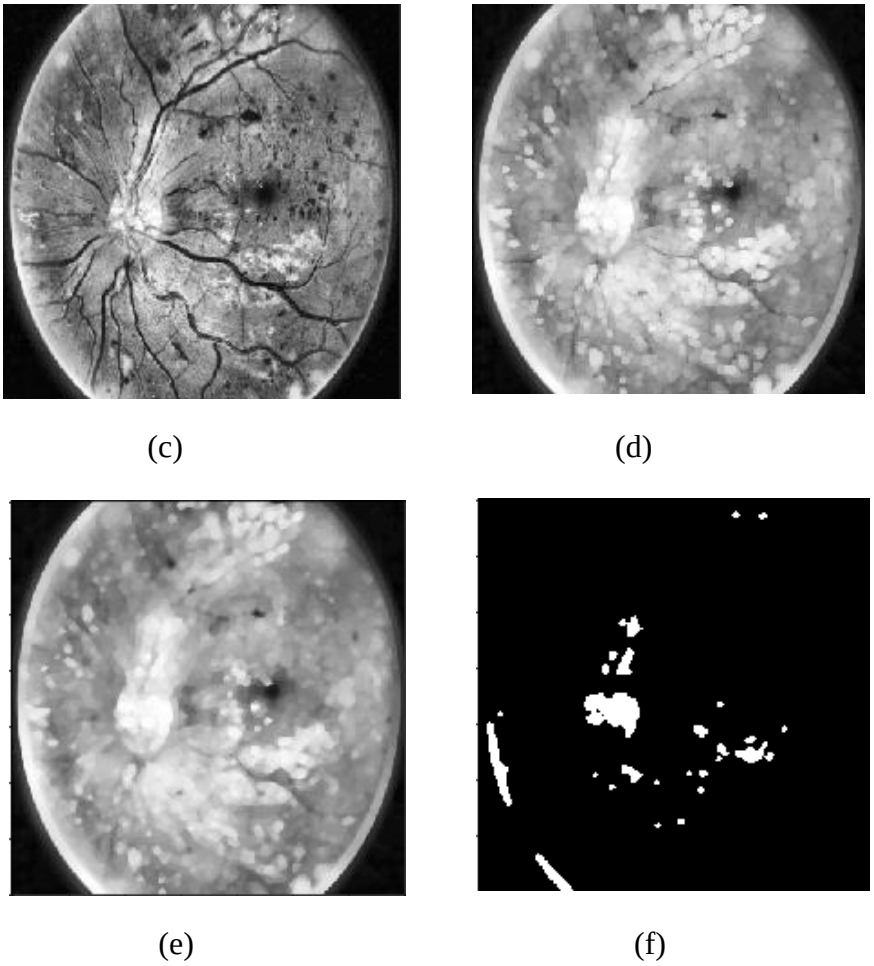
- **Exudate Detection:**

Using 6x6 ellipse shaped structuring element, morphological dilation is applied. Non linear median filter is used for noise removal. Exudates are in high intensity values. So it has been extracted using thresholding. After applying these preprocessing, pixels having intensity value higher than 235 are set to 255 and the rest of them are set to 0. Then traversing the image, area of exudates are calculated. The images of different steps are illustrated in Fig. 3.3.1.1.

Different stages of exudate preprocessing are illustrated in the following:



(a) (b)



(c) (d)

(e) (f)

Fig. 3.3.1.1. Steps of Exudate detection

- **Blood Vessel Detection:**

Blood vessel is one of the most important features for differentiating diabetic retinopathy stages. After obtaining the green channel image and improving the contrast of the image, several steps has been done for extracting blood vessel. Alternate sequential filtering (three times opening and closing) using three different sized and ellipse shaped structural element 5x5, 11x11 and 23x23 is applied on the image. Then the resultant image is subtracted from the input image. Subtracted image has lots of small noises. Those noises are removed through area parameter noise removal. Contours of each components including noises are found by using the function findContours( ) and calculate the contour area by using the function contourArea( ) and remove the noises having more than or equal to 200 pixels as area. Then the resultant image is binarized using a threshold value. Finally the

number of pixels that covers the blood vessels area are calculated. The images of different steps are illustrated in Fig. 3.3.1.2.

Different stages of blood vessel preprocessing are illustrated in the following:

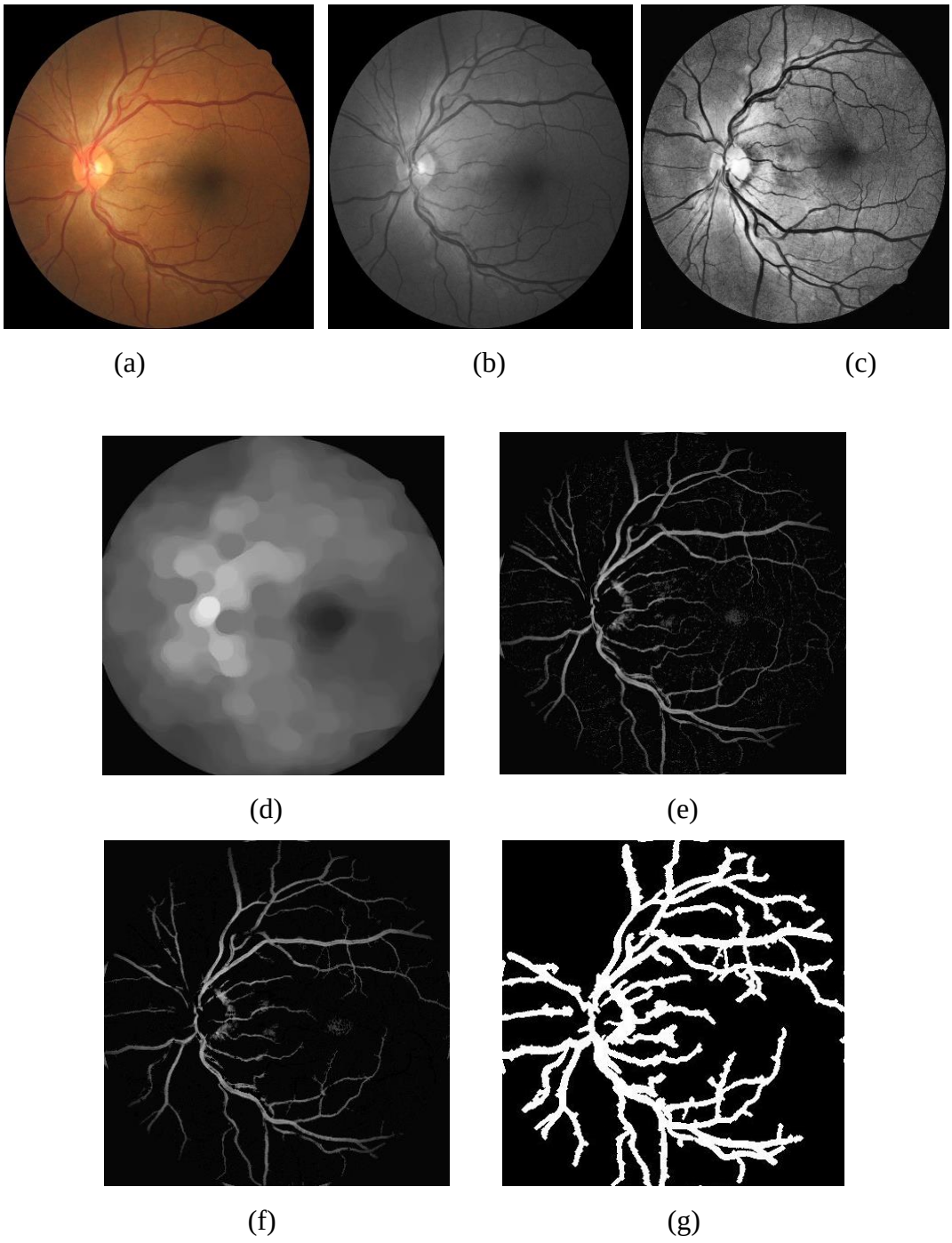


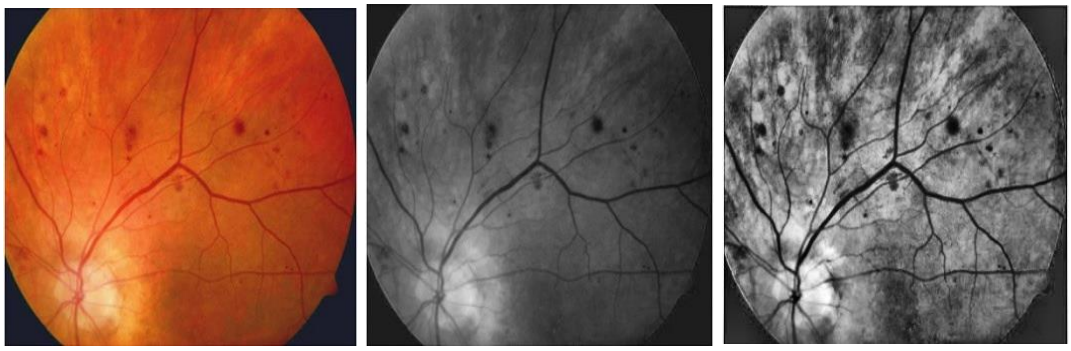
Fig. 3.3.1.2 Steps of Blood Vessel detection

- **Microaneurysm Detection:**

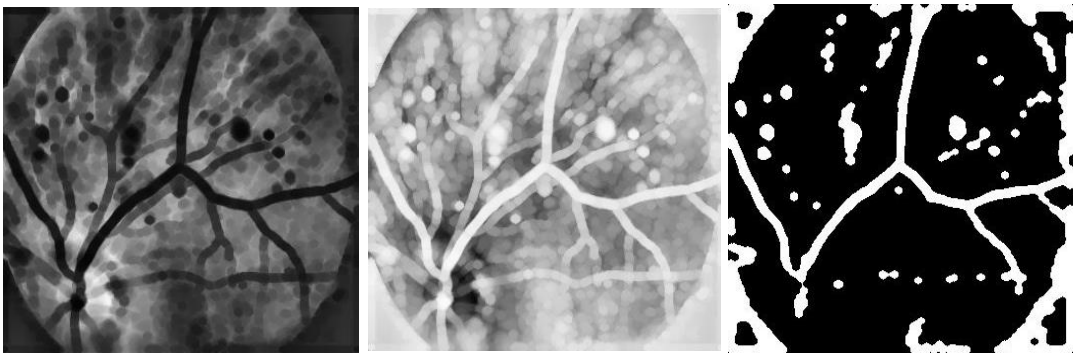
Green component of the RGB value is used to extract microaneurysm. For better contrast, CLAHE is used. Then median filter is used for noise removal. 7x7 ellipse shaped structural element is used for morphological operation. Morphological operation erosion is applied and then the image is inverted.

For joining the disjoint segments of blood vessel, morphological closing is used. Then the image has been binarized. As the blood vessel, haemorrhage and microaneurysm is having the almost same intensity, all these components will be detected altogether in the binarized image. Since microaneurysm is smaller in size, it has been extracted using contour area. The images of different steps are illustrated in Fig. 3.3.1.3.

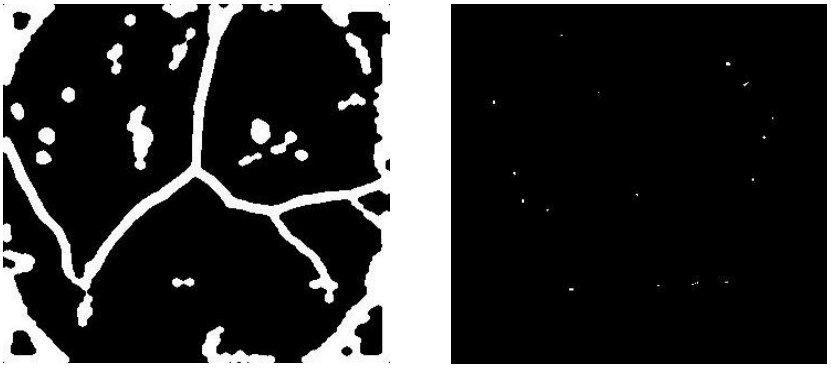
Different stages of microaneurysm preprocessing are illustrated in the following:



(a) (b) (c)



(d) (e) (f)



(g) (h)

Fig. 3.3.1.3. Steps of Microaneurysm detection

3.3.2 Feature Extraction:

Feature extraction will be done from preprocessed images shown in Fig. 3.3.1.1, 3.3.1.2 and 3.3.1.3. The features which are extracted to detect Diabetic Retinopathy are-

- Histogram of Exudates
- Zeroth hu moment of Exudates
- Histogram of Blood Vessels
- Zeroth hu moment of Blood Vessels
- Histogram of Microaneurysm
- Zeroth hu moment of Microaneurysm

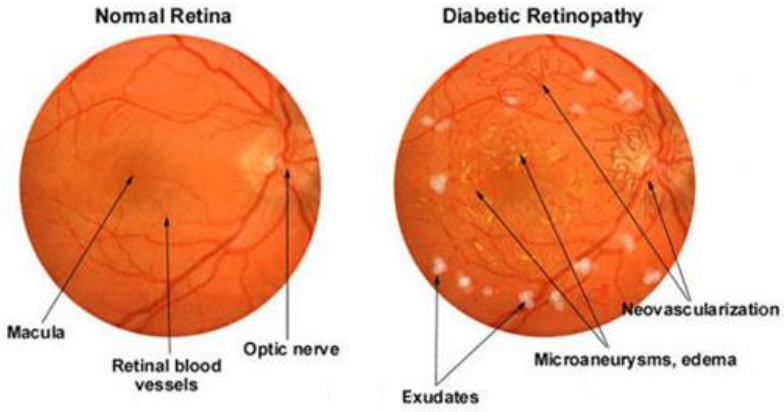


Fig. 3.3.2.1. Differences between a normal retina and DR affected retina

3.3.2.1 Histogram:

In an image processing context, the histogram of an image normally refers to a histogram of the pixel intensity values. This histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. For an 8-bit grayscale image, there are 256 different possible intensities, and so the histogram will graphically display 256 numbers showing the distribution of pixels amongst those grayscale values.

Histograms can also be taken of color images, either individual histograms of red, green and blue channels can be taken, or a 3-D histogram can be produced, with the three axes representing the red, blue and green channels, and brightness at each point representing the pixel count. The exact output from the operation depends upon the implementation, it may simply be a picture of the required histogram in a suitable image format, or it may be a data file of some sort representing the histogram statistics.

In this work, the white pixels are counted from the histogram to get the features from the binarized images of exudates, blood vessels and microaneurysm.

3.3.2.2 Hu Moments:

Central moments are translation invariant. But that is not enough for shape matching. We would like to calculate moments that are invariant to translation, scale, and rotation. We can in fact calculate such moments and they are called Hu Moments.

Hu Moments are a set of 7 numbers calculated using central moments that are invariant to image transformations. The first 6 moments have been proved to be invariant to translation, scale, and rotation, and reflection. While the 7th moment's sign changes for image reflection.

In this work, the zeroth moments from the binarized images of exudates, blood vessels and microaneurysm are used as features.

3.3.3 Classification:

Classification has been done using Random Forest classifier to classify the images into five classes as normal, mild NPDR, moderate NPDR, severe NPDR and PDR.

3.3.3.1 Classification using Random Forest:

Random Forest is an ensemble tree-based learning algorithm. The Random Forest Classifier is a set of decision trees from randomly selected subset of training set. It aggregates the votes from different decision trees to decide the final class of the test object. The fundamental concept behind random forest is a simple but powerful one - the wisdom of crowds. A large number of relatively uncorrelated trees operating as a committee will outperform any of the individual constituent models. Uncorrelated models can produce ensemble predictions that are more accurate than any of the individual predictions. The reason for this wonderful effect is that the trees protect each other from their individual errors as long as they do not constantly all errors in the same direction. While some trees may be wrong, many other trees will be right, so as a group the trees are able to move in the correct direction.

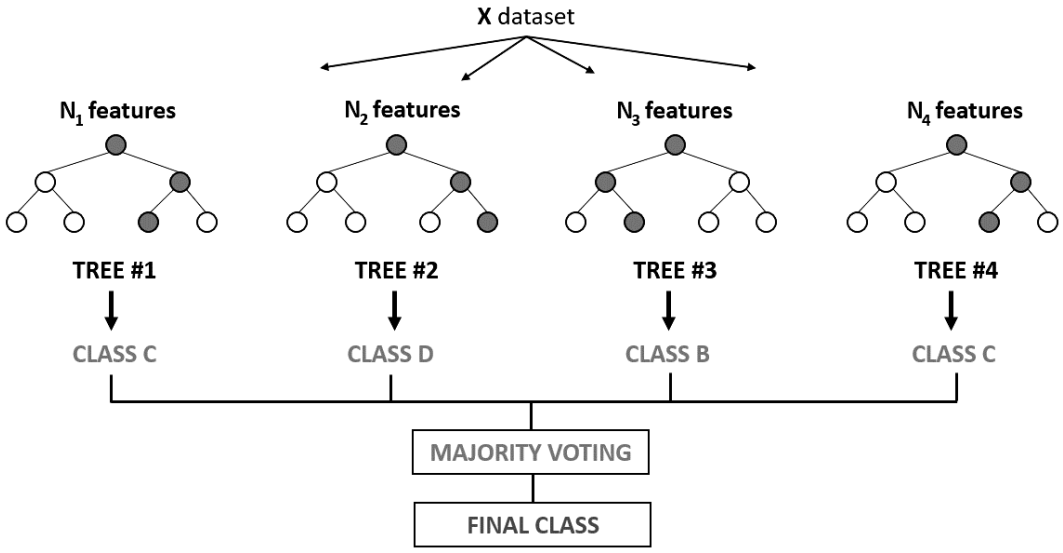


Fig. 3.3.3.1. Working procedure of a Random Forest classifier

Classifier parameters are fine tuned for the best performance. Since random forest generates decision trees for prediction, several parameters are responsible for performance analysis. Considering the features, we have used-

a) No. of estimators = Number of trees in the forest = 3000
b) Max_features = max number of features considered for splitting a node = Number of features
c) Criterion = Gini index

3.3.3.2 Classification using Support Vector Machine:

Support Vector Machine (SVM) classify the image into two classes such as Diabetic Retinopathy eye and healthy eye. As SVM is a binary classifier, our first task is to classify which eye is affected by Diabetic Retinopathy and which is a healthy one. After first classification, our next task is to use Support Vector Machine again. This time it is applied only on the affected ones. It will again classify which Diabetic Retinopathy is non-proliferative i.e. is in initial stage and which on is in proliferative i.e. is in severe state.

We have used Support Vector Machine because objective function in SVM is convex function which never stuck into the local maximum. Optimal hyperplane is the form of the separating hyperplane and objective function of optimization problem do not depend explicitly on dimensionality of the input vector but depends only on the inner products of two vectors. This fact allows constructing the separating hyperplanes in high dimensional spaces.

3.3.3.3 Classification using Naïve Bayes:

Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theorem. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e. every pair of features being classified is independent of each other.

Chapter 4

Experimental Studies

In our thesis work, we evaluate our work in terms of Accuracy, Sensitivity and Specificity. And the evaluation metrics are calculated and compared with several related works which are referenced.

4.1 Experimental Setup

To perform our work, Diabetic Retinopathy Classification from Retinal Images using Machine Learning Approach, we have used a computer with decent configuration as an experimental setup. For our thesis we have used following configuration as experimental setup-

- Intel Core i7 5th generation processor.
- 8GB RAM
- 1TB Hard-disc
- 2GB graphics card
- Jupyter Notebook
- Anaconda

4.2 Dataset

To evaluate our method, we have used a dataset named as Diabetic Retinopathy (Resized) from Kaggle. The dataset has a total of 13402 retinal images and corresponding levels of Diabetic Retinopathy for each image. The dataset has been split up to 75% as training data and 25% as test data. Therefore, 10052 training images has been used to train the model and it has been tested on 3350 images.

4.3 Evaluation

In this research, the summary of proposed system has been compared with evaluations of the existing methods. The Accuracy, Sensitivity, Specificity are

brought into play here as these have long used as important evaluation metrics in Diabetic Retinopathy classification.

Accuracy, Sensitivity and Specificity are computed as:

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total\ Number\ of\ Data} \tag{4.1}$$

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{4.2}$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} \tag{4.3}$$

4.4 Experimental Result and Analysis

For the efficiency judgment of the proposed method, experiments have been conducted on 3350 test images. After testing the images in this model using Random Forest classifier, confusion matrix of classification has been found. The confusion matrix has been shown in Table. 4.4.1.

Table 4.4.1. Confusion matrix for 5 classes

		Actual				
		Normal	Mild	Moderate	Severe	PDR
Predicted	Normal	989	100	114	63	66
	Mild	116	466	10	6	6
	Moderate	122	10	492	8	10
	Severe	82	2	12	316	8
	PDR	40	4	14	4	291

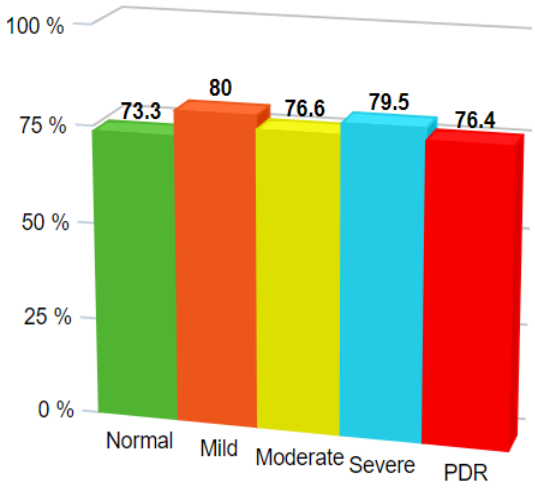


Fig. 4.4.1. Bar diagram of sensitivity for each class

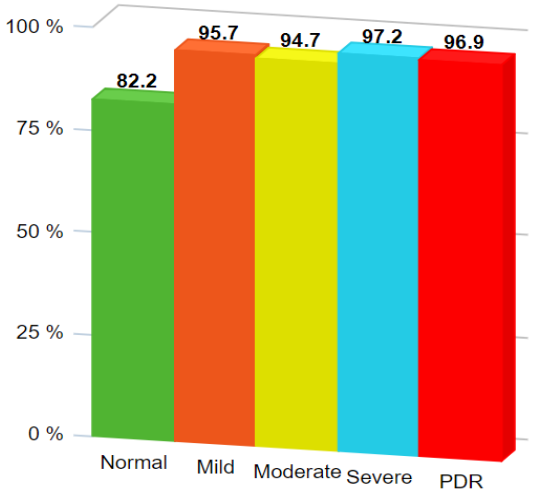


Fig. 4.4.2. Bar diagram of specificity for each class

Average sensitivity of the classification = **77.2%**

Average specificity of the classification = **93.3%**

Accuracy of the classification = **76.5%**

Since, the instances of each class are important in this work, we have calculated the evaluation metrics by using macro averaging method. Because macro average better reflects the statistics of the smaller classes, and so is more appropriate when performance on all the classes is equally important.

4.5 Result Comparison

In this work, finally the results of Random Forest classifier has been taken into account. Since Random Forest gives the best accuracy, sensitivity and specificity, Support Vector Machine and Naïve Bayes classifiers' results have been neglected. The results comparisons are shown in table 4.5.1.

Table 4.5.1. Comparison of results of different classifiers

Sl. No.	Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
1	Random Forest	76.5	77.2	93.3
2	Support Vector Machine	47.4	48.1	61.3
3	Naïve Bayes	36.8	43	58.1

4.6 Discussion on Results

The evaluation measures are compared with the related works in terms of number of classes, methods used, features etc.

Table 4.6.1. Comparison of Automatic detection of the DR stages by various researchers

Sl. No.	Reference	Number of Classes	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
1	Sinthanayothin et al. [6]	2	Moat operator	Not reported	80.2	70.6
2	Singalavanija et al. [16]	2	Exudates, Blood vessel, Microaneurysm	Not reported	74.8	82.7
3	Kahai et al. [17]	2	Decision support system	Not reported	100	63
4	Wang et al. [15]	4	Area of blood vessel	84	91.7	100

Sl. No.	Reference	Number of Classes	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
5	Nayak et al. [5]	3	Blood vessel, exudates, texture	93.6	90.3	100
6	Acharya et al. [18]	5	Higher order spectra	82	82.5	88.9
7	Lim et al. [19]	5	Blood vessels, exudates, microaneurysms, haemorrhages	85.9	82	86
8	This work	5	Histogram of blood vessel, exudates, microaneurysm, Zeroth hu moment of blood vessel, exudates, microaneurysm	76.5	77.2	93.3

DR was differentiated from a normal retina using image processing algorithms by Sinthanayothin et al. [6]. In their method, retinal images were preprocessed using adaptive local contrast enhancement. Their system, based on a multi-layer perceptron neural network, yielded a sensitivity of 80.21 percent and a specificity of 70.66 percent.

A decision support system for the early detection of the DR was proposed by Kahai et al. [17]. The detection rule is based on a binary-hypothesis testing problem with yes–no decisions. The Bayes optimality criterion was applied to fundus images for the early detection of the DR. This system was able to identify the presence of microaneurysms with a sensitivity of 100 percent and specificity of 67 percent accurately.

Wang et al. [15] have classified normal, mild DR, moderate DR, severe DR, and PDR stages using morphological image-processing techniques and a feed forward neural network [18]. In their work, the area and perimeter of the RGB components of the blood vessels are chosen as the features for the classifier. The average

classification efficiency of their system was 84 percent, the sensitivity was 90 percent, and the specificity was 100 percent. Using the area of the exudates, blood vessels, and texture parameters, the fundus images were classified into normal, NPDR, and PDR [4]. They demonstrated a classification accuracy of 93 percent, a sensitivity of 90 percent, and a specificity of 100 percent.

Recently, Acharya et al. [18] automatically identified normal, mild DR, moderate DR, severe DR, and PDR stages using the bispectral invariant features of higher-order spectra techniques and a SVM classifier [18]. They obtained an average accuracy of 82 percent in identifying the unknown class, a sensitivity of 82 percent, and a specificity of 88 percent.

In this work, the DR classes are classified into five classes using exudates, microaneurysms, and blood vessel areas. Random Forest is used for the classifier. The classifier is able to identify the unknown class accurately with an efficiency of more than 76.5 percent with sensitivity 77.2 percent and specificity 93.3 percent.

Chapter 5

Conclusion

After studying the existing systems, we conclude that our proposed technique is successfully detecting Diabetic Retinopathy. Along with this, the proposed method is classifying into five classes of Diabetic Retinopathy.

5.1 Findings

Classification has been done based on three features- area of exudates, area of blood vessel and area of microaneurysm. And using this features, we have classified into five classes as normal eye, mild Non Proliferative DR, moderate Non Proliferative DR, severe Non Proliferative DR and Proliferative DR. Using Random Forest classifier, we have gained accuracy=76.5%, sensitivity=77.2% and specificity=93.3%. The metrics we have found in this work are compared with the existing works.

5.2 Future Work

In this thesis work, we have performed the Diabetic Retinopathy classification using Random Forest classifier with some essential features like exudates, blood vessel and microaneurysm. In future, we hope to make it work for some more classifiers like K-Nearest Neighbor classifiers and so on using some secondary features like haemorrhage also. And we can perform this classification method using larger dataset of infected eyes using neural network model in future.

References

[1] Akara Sopharak, Matthew N. Dailey, Bunyarit Uyyanonvara, Sarah Barman, Tom Williamson, Khine Thet Nwe& Yin Aye Moe (2010).“Machine learning approach to automatic exudate detection in retinal images from diabetic patients.”*Journal of Modern Optics, 57:2, 124-135.*

[2] Shailesh Kumar and Basant Kumar (2018). “Diabetic Retinopathy Detection by Extracting Area and Number of Microaneurysm from Colour Fundus Image.” *5th International Conference on Signal Processing and Integrated Networks (SPIN) (2018):, 359-364.*

[3] Sohini Roychowdhury, Dara D. Koozekananiand Keshab K. Parhi (2014). “DREAM: Diabetic Retinopathy Analysis Using Machine Learning.” *IEEE Journal of Biomedical and Health Informatics (Volume: 18 , Issue: 5 , Sept. 2014), 1717 – 1728.*

[4] Liu, Z.; Chutatape, O.; Krishna, S.M. Presented at the *19th IEEE Conference on Engineering in Medicine and Biology Society, Chicago, USA, Oct 30–Nov 2, 1997.*

[5] Nayak, J., Bhat, P. S., Acharya, U. R., Lim, C. M., and Kagathi, *M. Automated identification of different stages of diabetic retinopathy using digital fundus images.*

[6] Sinthanayothin, C.; Boyce, J.F.; Williamson, T.H.; Cook, H.L.; Mensah, E.; Lal, S.; Usher, D. J. *Diabet. Med. 2002, 19, 105–112.*

[7] Usher, D.; Dumskyj, M.; Himaga, M.; Williamson, T.H.; Nussey, S.; Boyce, J. J. Diabet. Med. 2004, 21, 84–90.

[8] Kavitha, D.; Shenbaga, S.D. Presented at the *2nd ICISIP Conference on Intelligent Sensing and Information Processing, Madras, India, Jan 4–7, 2005.*

[9] Osareh, A.; Mirmehdi, M.; Thomas, B.; Markham, R.*In Medical Image Understanding and Analysis Conference;Claridge, E., Bamber, J., Eds.; BMVC Press: UK, 2001;pp 49–52.*

[10] Osareh, A.; Mirmehdi, M.; Thomas, B.T.; Markham, R. In MICCAI'02: Proceedings of the *5th International Conference on Medical Image Computing and Computer-Assisted Intervention-Part II; Springer- Verlag: London, UK, 2002; pp 413–420.*

[11] Zhang, X.; Chutatape, O. Presented at the *IEEE Conference on Computer Vision and Pattern Recognition, San Diego, USA, June 20–25, 2005.*

[12] Katarzyna, S.; Adam, S.; Radim, C.; Georg, M. Segmentation of Fundus Eye Images Using Methods of Mathematical Morphology for Glaucoma Diagnosis, *Lecture Notes in Computer Science, 2004.*

[13] Sanchez, C.I.; Hornero, R.; Lopez, M.I.; Poza, J. Presented at the *26th IEEE Conference on Engineering in Medicine and Biology Society, California, USA,Sep 1–4, 2004.*

[14] Walter, T.; Klevin, J.C.; Massin, P.; Erginay, A. *IEEE Trans. Med. Imaging 2002, 21, 1236–1243.*

[15] Wang, H.; Hsu, W.; Goh, K.G.; Lee, An Effective Approach to Detect Lesions in Color Retinal Images, M.L. Presented at the *IEEE Conference on Computer Vision and Pattern Recognition, South Carolina, USA, June 13–15, 2000.*

[16] Singalavanija, A., Supokavej, J., Bamroongsuk, P.Sinthanayothin, C., Phoojaruenchanachai, S., and Kongbunkiat, V. *Feasibility study on computeraided screening for diabetic retinopathy. Jap. J.Ophthalmology*, 2006, 50(4), 361–366.

[17] Kahai, P., Namuduri, K. R., and Thompson, H. A decision support framework for automated screening of diabetic retinopathy. *Int. J. Biomed. Imaging, 2006, 1–8.*

[18] Acharya, U. R., Chua, K. C., Ng, E. Y. K., Wei, W., and Chee, C. Application of higher order spectra for the identification of diabetes retinopathy stages. *J. Med. Systems, 2008, 32(6), 481–488.*

[19] C M Lim, U R Acharya, E Y K Ng, C Chee and T Tamura, *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 2009 223: 545

[20] Ege, B.; Hejlesen, O.; Larsen, O.; Moller, K.; Jennings, B.; Kerr, D.; Cavan, D.A. *Computer Methodoly Programs Biomed. 2000, 62, 165–175.*

[21] L. Zhang, Q. Li, J. You and D. Zhang, "*A Modified Matched Filter With Double-Sided Thresholding for Screening Proliferative Diabetic Retinopathy*," in IEEE Transactions on Information Technology in Biomedicine, vol. 13, no. 4, pp. 528-534, July 2009.

[22] M. Gandhi and R. Dhanasekaran, "*Diagnosis of diabetic retinopathy using morphological process and SVM classifier*", 2013 International Conference on Communication and Signal Processing, Melmaruvathur, 2013, pp. 873-877.