Analysis of Diabetic Retinopathy from the Features of Color Fundus Images using Classifiers

Gandhimathi. K[1], Ponmathi. M[2], Arulaalan. M[3] and Samundeeswari. P[4]

[1,2]Assitant Professor/ CSE
[3]Associate Professor/ ECE
[4]Assitant Professor/ ECE
[1,2]Idhaya Engineering College for Women, Chinnasalem, Tamilnadu, India
[3]Alpha college of Engineering & Technology, Pondicherry, India
[4]Alpha college of Engineering & Technology, Pondicherry, India

Abstract— **Diabetic Retinopathy (RD) is a complication of Diabetes that is caused by changes in the blood vessels of the retina. The development of abnormal vessels in the retina may lead to vision blur, distortion or even may cause visual impairment. Hence early detection of RD is essential to avoid blidness. In this paper, an automatic RD detection system has been developed to analyze the presence of lesion (presence of new cells in the retina) in fundus images using the two different classifiers – Multi-Layer Perceptron (MLP) and Support Vector Machine (SVM). The fundus images have been preprocessed to improve the performance of early detection by performing filter, watershed transformation segmentation method and feature extraction method. The vessels like candidate segments can be detected based on the different parameters associated with shape, contrast, position, orientation, brightness and the line density. The proposed automatic classifier using SVM performs better detection of the abnormal in the fundus image.**

Keywords— *Diabetic Retinopathy (DR), Optic Disk (OD), Multi-Layer Perceptron (MLP), Watershed, Support Vector Machine (SVM) and Micro Aneurysms (MAs).*

I. INTRODUCTION

DIABETIC retinopathy is an important complication of diabetes and a leading cause of blindness. The occurrence of the DR damages the tiny blood vessels inside the retina and light-sensitive tissue at the back of the eye. Early detection of the disease via regular screening is particularly important to prevent vision loss. The importance of detecting Micro Aneurysms (MAs) is underscored by the fact that they are the first clinically evident sign of diabetic non-proliferative eye disease. Hence the recognition of MAs can be the first step in prevention of diabetic retinopathy progression to the proliferative stage and consequent severe visual loss. MA's is one of the earliest clinical signs of development of diabetic retinopathy. They generally appear as small round red spots and their diameters are generally less than the diameter of main blood vessels. Increase in number of MAs leads to progression of retinopathy.

Exudates are one of the most important primary features of diabetic retinopathy, which are responsible for hazy views and blindness. It appears like as yellow flecks and caused by lipid occlusions from the damaged blood vessels. Retinopathy is a progressive disease that can advance from mild stage to Proliferative stage. It has three stages: (i) early stage, Non-Proliferate Diabetic Retinopathy (NPDR) or background retinopathy, (ii) maculopathy and (iii) progressive or proliferate retinopathy. These stages of DR are shown in Fig. 1.
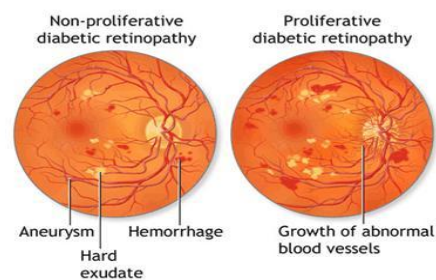

Fig. 1 Main Stages of Retinopathy with the Disorders

The early stage is further classified as mild NPDR and moderate to severe NPDR. This paper merely deals with detection of mild NPDR. In mild NPDR, signs such as MAs, hard or intra-retinal exudates, dot and blot hemorrhages are seen in the retinal images. MAs are first detectable signs of retinopathy which appears like small, round, dark red dots with sharp margins and are often temporal to macula. Their size ranges from 20 to 200 microns (i.e., less than $1/12^{th}$ the diameter of an average optic disc). Hard exudates are shiny, irregularly shaped and found near prominent MAs or at the edges of retinal edema. In the early stage, the vision is rarely affected and the disease can be identified only by regular dilated eye examinations.

In this paper, an automatic detection system is proposed with an algorithm that can detect the abnormal blood vessels by comparing with normal retinal images. This process has been done by three steps. Initially, the fundus image has been preprocessed and segmented with the Watershed transformation method. The result of the preprocessed image will be taken for feature selection based on different parameters. These parameters of the normal and abnormal retinal images are trained with two classifiers namely MLP and SVM. This overall processes of detecting abnormality in the retinal image as shown in the Fig.4.

II.   METHODOLOGY

A.   Fundus Image Database

Database containing 50 fundus images of retina have been collected from MV Hospital for Diabetes & Research Centre, Chennai. This database contains both normal and abnormal fundus angiographic retinal images. This database has been used during testing process of the trained classifiers and to analysis the performance of both classifiers (MLP and SVM).

For training process of classifiers, the publicly available database of fundus angiographic retinal images was downloaded from the web site which also contains (http://www.it.lut.fi/project/imageret) both normal and abnormal images. It contains 120 retinal images captured by fundus camera. In this database 80 images are affected by diabetic retinopathy and the rest of the images are normal retinal images.

B.   Preprocessing

In detecting abnormalities associated with fundus image the images have to be pre-processed in order to correct the problems of uneven illumination, non-sufficient contrast among microaneurysms, image background pixels and presence of noise in the fundus image. Aside from before mentioned problems, this section is also responsible for colour space conversion and image size standardization to improve the efficient performance of the classifier techniques. The pre-processing method of the fundus images can be implemented by using filtering, segmentation and feature extraction methods.

1.   Filtering

In fundus image, the blood vessels have lower reflectance than other retinal surfaces; hence they appear darker relative to the background. The Gaussian filter uses a Gaussian shaped curve to approximated the typical vessel cross-sectional gray-level profile (i.e. outermost vessel pixels are brighter than the inner vessel pixels), but some of the blood vessels include a light streak which runs down the central length of the blood vessel. So initially, the green plane of the fundus image has been obtained to remove this brighter strip. Then it is filtered by applying a Gaussian filter. Fig 2 shows the green plane of fundus image and fundus image after Gaussian filtered. The green channel shows the background gray levels are higher than that of vessels gray levels. The inverted image was filtered with a 2-D Gaussian function (with a standard deviation equal to 2 pixels) to prevent over-segmentation (Fig 2).
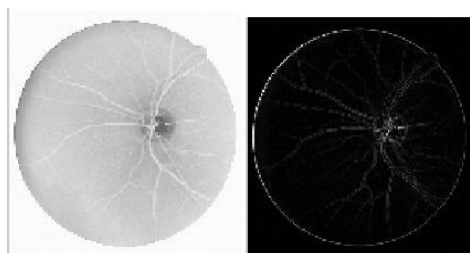

Fig. 2 Filtered with green plane (left), After Gaussian Filtered fundus image (right).

2.   Segmentation

The watershed transform using Meyer's algorithm has been adapted to segment the blood vessels and microaneurysms from the Gaussian filtered fundus image. After watershed transformation, this image has been subjected to thinning process that makes easier identification of the retinal vessels. The thinning process is a morphological operation that has been used to remove selected foreground pixels from binary images. After that the image has been modified to remove variations in background intensity by enhancing the image, and then the MAs were extracted from this image. Each MAs candidate is then classified according to its intensity and size by the applying a set of rules derived from a training set of images. Contrast normalization is also obtained by using the watershed transform to derive a region that contains no vessels or other lesions. Dots within the vessels are handled successfully using a local vessel detection technique.

(i) After thinning   (ii) Edge Detection

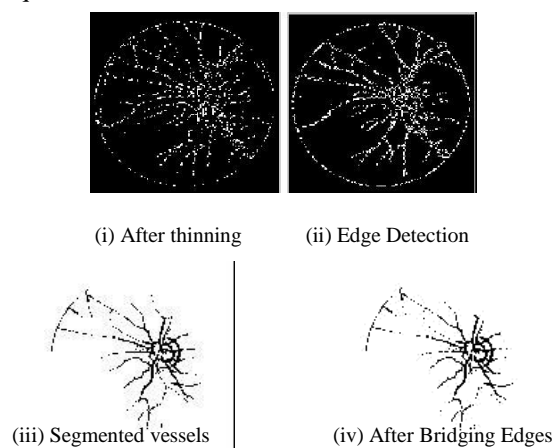(iii) Segmented vessels   (iv) After Bridging Edges

Fig. 3 Different Stages of Segmentation

In [13] and [14], vessel detection is achieved by intensity based region growing from the candidate center. It results for detection of individual MAs and for the detection of images containing MAs. In [15] and [17], an independent vessel detection method is first employed so that MA candidates can be rejected where vessel detection was successful. Images containing MAs are to be detected with sensitivity 85.4% and specificity 83.1%. The dark ridges formed by the vessel center lines may be detected using the ridge strength (contour curvature), K given by

$$K = \frac{L_x^2 L_{yy} + L_y^2 L_{xx} - 2L_x L_y L_{xy}}{(L_x^2 + L_y^2)^{3/2}} \qquad (1)$$

The watershed regions are calculated using Meyer's algorithm [2], implemented in the MATLAB image processing toolbox (Mathworks Inc.). The watershed transform was also used by Walter et al. to detect vessels across the retina [1]. The binary image of the watershed lines is thinned (Fig 3(i)), such that only pixels at vessel bifurcations have more than two neighbors

3.   Feature Extraction

The abnormal vessels that can be identified and measured based on the 10 feature parameters. The feature parameters have been characterized based on their position, shape, orientation, brightness and line density. After the thinning process of the image (Fig. 3(i)), the angle was then calculated

from the gradient of the fitted line and constrained to lie in the right-hand side of the plane. Firstly, a 19X19 pixel median filter was applied to remove smaller vessels. Next a threshold was applied to select the darkest 20% of pixels, which were assumed to make the pixels that belong to the major blood vessels. The centroid of the result was taken as the approximate origin of the major vessels. To train the classifiers the following mentioned features are extracted which are explained below

(i) Length of Segment: The length of the segment in pixels.

(ii) Gradient: The mean gradient magnitude along the segment using the Sobel gradient operator.

(iii) Direction: The angle between a tangent to the segment center point and a line from its center point to the vessel origin. The feature is based on the observation that normal vessels tend to radiate from the vessel origin towards the edge of the disc, whereas the direction of new vessels is more random.

(iv) Tortuosity: The sum of the absolute changes in the tangential direction along segment path given below,

$$T = \frac{1}{n-1}\sum_{i=1}^{n-1}|\theta_{i+1} - \theta_i| \qquad (2)$$

(v) Grey Level: The normalized mean segment grey level (g_{norm}) where g_i is the grey level of the ith segment pixel, G_{max} is the maximum level values and G_{min} is the minimum grey level values in the original image, respectively.

$$g_{norm} = \frac{1}{G_{max} - G_{min}}\left[\left(\left(\frac{1}{n}\right)\sum_{i=1}^{n}g_i\right) - G_{min}\right] \quad (3)$$

(vi) Distance from Vessel Origin: The distance (in pixels) from the center of the segment to the disc vessel origin. This feature was included to test for any positional dependency of the segment within the disc.

(vii) Number of Segments: The total number of segments following candidate segmentation. Images with new vessels tend to have a higher number of segments.

(viii) Mean Vessel Width: An edge map was generated using the canny edge detector Fig. 3(ii). The distance from each segment point to the closest edge point is assumed to be the vessel half-width at that point. If the distance is too large it is assumed that the true edge has not been identified, probably due to a very fine vessel, and a width of zero is used.

(ix) Grey Level Coefficient of Variation: This measure was based on the observation that new vessels appear less homogeneous than normal vessels. Grey Level Coefficient of Variation is calculated as the ratio of the mean and standard deviation of the segment grey level values.

c. Classification and Training

The feature parameters that have been calculated from the images have been taken as the training dataset. The training dataset which contains the several features of each image forms the databases.

To training these features and also to classify the abnormalities from each segments of the image, the two different classifiers such as Multi-Layer Perceptron (MLP) and Support Vector Machine (SVM).
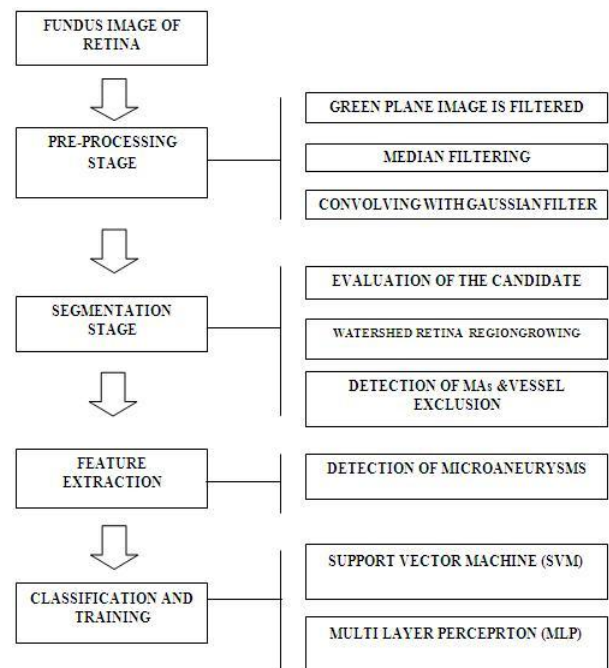

Fig. 4: Flow Diagram of Proposed System

1. Multi-Layer Perceptron classifier

The MLPs are an important class of NNs that can represent nonlinear functional mappings between a set of input variables and a set of output variables [10]. A MLP with enough units in a single hidden layer can approximate any function, provided the activation function of the neurons satisfies some general constraints [9, 10]. From these considerations, we decided to use a MLP with one hidden layer and the optimum number of hidden neurons was experimentally determined. For neuron activation function in the hidden layer, we chose the hyperbolic tangent sigmoid function (tan-sigmoid), an antisymmetric function in the interval (-1, 1). Tan-sigmoid satisfies the constraints in [11] and [12]. Moreover, it improves the learning speed of MLP [9]. In the output layer uses logistic sigmoid activation function, that also satisfies the aforementioned constraints and whose outputs lie in the range (0, 1). This choice was motivated by the fact of interpreting the outputs of the network as posterior probabilities [5].

The problem of training a NN can be formulated in terms of the minimization of an error function. The choice of a suitable error function and minimization algorithm can improve the performance of MLP. It has been demonstrated [5] that a cross-entropy error function simplifies the optimization process when the logistic sigmoid activation function is used in the output layer. Therefore, we considered this function as an appropriate choice in our study.

Regarding the minimization algorithm, numerous choices are available for MLP classifier. We selected the 13 scaled conjugate gradients algorithm which provides guarantee to

the error function; it not increases during training [5]. Moreover, it generally shows faster convergence when compared to gradient descent-based techniques or even conventional conjugate gradient algorithms.

2. *Support Vector Machine*

The original SVM algorithm is a linear classifier which finds the best hyperplane separation of two classes. However, a kernel function can be used to transform the features to a higher dimensional space. Although the SVM finds a linear hyperplane in the transformed space, the chosen hyperplane is likely to be nonlinear in the original feature space. The kernel function ($K(x_i, x_j)$) used here was a radial basis function is given by,

$$K(x_i, x_j) = \exp(-\gamma \|X_i - X_j\|^2) \gamma > 0 \qquad (4)$$

Where x_i and x_j are the feature vectors for the two classes and γ is a configurable parameter. In addition to that, cost or penalty function weight (C) is also configurable. All the features were normalized before classification by using below expression,

$$\bar{f} = \frac{f - m}{s} \qquad (5)$$

Where f is the feature value to normalize, $\bar{f}$ is the normalized value, m is the mean and s is the standard deviation for the feature.

The SVM estimates a probability of abnormality for each vessel segment [9]. For the detection of abnormal images the single segment with the highest abnormality probability was selected and compared with a threshold. Different operating points may be chosen by varying the abnormality score threshold.

Both the Classifiers shown the better result by training the datasets with different features but comparatively SVM made impact that it is better than MLP. The performance was assessed in terms of an image-based criterion and a lesion-based criterion (pixel resolution) [3]. The image-based criterion accounted for the ability of the algorithm to separate pathological images from normal ones on the basis of the presence or absence of MAs. With a lesion-based criterion, we examined the number of MAs in the images that were correctly detected. *f. Detection and E* detect only microaneurysms and remove all the false positives introduced in the earlier stages, by using a Boolean operation we combine the two images obtain feature based AND operation. In feature based AND, the ON pixels in one binary image are used to select objects in another image [11].

Here we use the image with objects with sharp edges to select object in the image with red spots, because in the last one the lesions are detected completely, not only their contours. But also in this way, we obtain lesions characterized by the two desired features: yellowish color and sharp edge.

After that, we also get some false positives due to the papillary region and some artifacts near the vessels. To reduce them, we remove a dilated version of the segmentation, result of the detection of the optic disk. The sensitivity and the specificity is calculated for the resulting image. Sensitivity (5) means the percentage of abnormal fundus classified as abnormal by the procedure. Specificity (6) means the percentage of normal fundus classified as normal by the procedure. The higher value of sensitivity and specificity provides the better classificaion. Sensitivity and specificity values [5] can be calculated as follows.

$$\text{Sensitivity} = \frac{T_p}{T_p + F_n} \qquad (6)$$

$$\text{Specificity} = \frac{T_n}{T_n + F_p} \qquad (7)$$

Where FP, TP, FN, TN mean False Positives, True Positives, False Negatives, True Negatives respectively.

A screened fundus is considered as a true positive if the fundus is really abnormal and if the screening procedure also classified it as abnormal. Likewise, a true negative means that the fundus is really normal and the procedure also classified it as normal. The false positive means that the fundus is really normal, but the procedure classified it as abnormal. The false negative means that the procedure classified the screened fundus as normal, but it really is abnormal retinal fundus image. In this stage we detect the microaneurysms in the fundus retinal image and classify diabetic retinopathy.

III. RESULT

The trained network is tested with 50 samples of fundus image with the presence of DR. The network shows good performance detecting 45 samples accurately. The accuracy of the classifier is measures in terms of Recall and Precision. Recall marks the ability of the algorithm to detect whether the image is found with abnormal vessels or not and precision marks the success of the classifier at excluding normal vessels which are not infected. The performance of the classifiers are verified in terms of False Positive (FP), False Negative (FN), True Negative (TN) and True Positive (TP) are given as follows:

- True Positive (TP): An image predicted to be in class C_j, and is actually in it.
- False Positive (FP): An image predicted to be in class C_j, but actually not in it.
- True Negative (TN): An image not predicted to be in class C_j, but is actually not in it.
- False Negative (FN): An image not predicted to be in class C_j, but is actually in it.

From the above classes we calculate the values of precision and recall as follows:

$$Accuracy = \frac{TP + TN}{TP + TP + FP + FN} \qquad (8)$$

The system has been tested with an image database which is obtain from the MV Hospital, that has different normal and abnormal fundus images. Table 1 show that the identification of the blood vessels at various instance (images) and the result analysis.

TABLE 1 VESSEL DETECTION ANALYSIS

IMAGE INDEX	VESSEL DETECTED	VESSELS CONSIDERED
IMG 1	20	6
IMG 2	18	8
IMG 3	28	4
IMG 4	16	10
IMG 5	30	8
IMG 6	12	6
IMG 7	20	15
IMG 8	15	7
IMG 9	26	9
IMG 10	17	5

From the image database, the images with the presence of abnormal vessels in the fundus image are taken for the training and detection of DR. The above table 1 lists out the number of total number of detected vessels and the number of vessels that has the abnormalities. The system after completing the clustering process the blood vessels the same group of blood vessels that find the difference with the normal vessels considered as abnormal vessels and taken for training dataset.

The propose system which overcomes the manual grading found to be more accurate. From the training and analysis the accuracy of the classifier is calculated. Table 2 shows the confusion matrix for the test samples. From which the classification accuracy is calculated.

TABLE 2 CONFUSION MATRIX FOR SVM

ACTUAL / PREDICTION	INFECTED	NOT INFECTED
INFECTED	40 (TP)	3 (TN)
NOT INFECTED	2 (FP)	35 (FN)

The above Table 2 figures out the confusion matrix that illustrates the classification accuracy of the classifier.

Accuracy of the proposed system over SVM as follows:

$$Accuracy = \frac{TP+TN}{TP+TP+FP+FN} = \frac{35+40}{35+40+2+3} = 0.9375 / 100$$
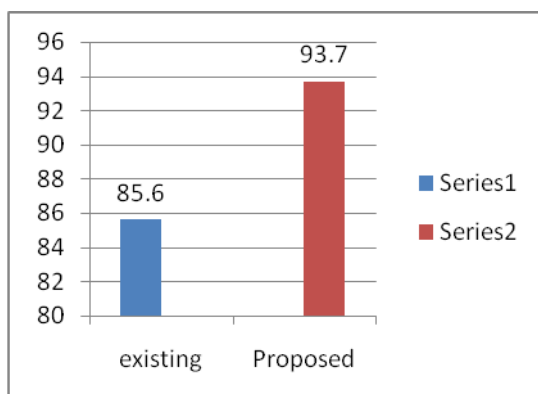
$$= 93.7\% \quad (6)$$

Fig. 5 Comparison of Classification Accuracy

The accuracy of the proposed classifier system (using SVM) has been achieved 93.7%. Fig. 5 illustrates the comparison of the classification accuracy of proposed system against the existing system (MLT). Therefore, the proposed system is more robust and can be used by the ophthalmologist in the detection of Diabetic Retinopathy in an efficient way.

IV. CONCLUSION AND FUTURE WORK

A system of automatic detection for Diabetic Retinopathyis has been proposed to perform early detect of the abnormal vessels present in the Fundus image and to make convenient for the Ophthalmologist to examine the patient's eye in an efficient way. The proposed classifier method detects only mild Non-Proliferative Diabetic Retinopathy (NPDR) symptoms. As a future work we can extend this work to detect moderate and severe NPDR symptoms, cotton wool spots, venous beading, venous loops, Intra-retinal micro vascular abnormalities (IRMA) and also classify the grade of severity about the abnormality condition. This automated system of detecting the abnormal vessels can be very helpful to reduce the manual grading of the visual impairment processing techniques. This method of early detection is very much useful to avoid cause of blindness problem.

REFERENCES

[1] Keith A. Goatman and Alan D. Fleming, "Detection of New Vessels on the Optic Disc Using Retinal Photographs", *IEEE Trans. Med. Imag*, vol. 30, no. 4, pp. 927-979. ARPIL 2011.

[2] Meindert Niemeijer and Bram van Ginneken, "Automatic Detection of Red Lesions in Digital Color Fundus Photographs", *IEEE Trans. Med. Imag*, Vol. 24, no. 5, pp: 584-592, MAY 2005.

[3] Meindert Niemeijer and Bram van Ginneken, "Retinopathy Online Challenge: Automatic Detection of Microaneurysms in Digital Color Fundus Photographs", *IEEE Trans. Med. Imag*.

[4] Alan D. Fleming,Keith and A. Goatman, "Automated Assessment of Diabetic Retinal Image Quality Based on Clarity and Field Definition", *IOVS*, Vol. 47, no. 3, pp: 1120-1126, March 2006.

[5] R.Priya and P. Aruna, "SVM and Neural Network based Diagnosis of Diabetic Retinopathy", *International Journal of Computer Applications* (0975 – 8887) Vol. 41– no.1, March 2012.

[6] Joes Staal, "Ridge-Based Vessel Segmentation in Color Images of the Retina", *IEEE Trans. Med. Imag*, Vol. 23, No. 4, pp: 501-509, APRIL 2004.

[7] Meindert Niemeijer," Retinopathy Online Challenge: Automatic Detection of Microaneurysms in Digital Color Fundus Photographs", *IEEE Trans. Med. Imag*, Vol. 29, no. 1, JANUARY 2010.

[8] M. Usman Akram and Shehzad Khalid ," Identification and classification of microaneurysms for early detection of diabetic retinopathy", *ELSEVIER, Pattern Recognition*46(2013)107–116.

[9] Bernhard M. Ege and Ole K. Hejlesen, "Screening for diabetic retinopathy using computer based image analysis and statistical classification", *ELSEVIER,Computer Methods and Programs in Biomedicine* 62 (2000) 165–175.

[10] María García and Clara I. Sánchez, "Neural network based detection of hard exudates in retinal images", *ELSEVIER, computer methods and programs in biomedicine* 93(2009)9–19.

[11] Cemal Kosea and Ugur S¸ evikb, "Simple methods for segmentation and measurement of diabetic retinopathy lesions in retinal fundus images", *ELSEVIER, computer methods and programs in biomedicine* 107(2012) 274–293.

[12] R.J. Winder and P.J.Morrow, "Algorithms for digital image processing in diabetic retinopathy", *ELSEVIER, Computerized Medical Imaging and Graphics* 33 (2009) 608–622.

[13] T. Spencer, J. A. Olson, K. C. McHardy, P. Sharp, and J. V. Forrester,"Image-processing strategy for the segmentation and quantification of micro aneurysms in fluorescein angiograms of the ocular fundus," *Comput. Biomed. Res.,* vol. 29, no. 4, pp. 284–302, 1996.

[14] M. J. Cree, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "A fully automated comparative micro aneurysm digital

detection system," *Eye*, vol. 11, no. 5, pp. 622–628, 1997.

[15] T. Walter and J. C. Klein, "Automatic detection of micro aneurysms in Color fundus images of the human retina by means of the bounding box closing", *In Proceedings of Medical Data Analysis. London,* U.K.: Springer-Verlag, 2002, vol. 2526, Lecture Notes in Computer Science, pp. 210–220.

[16] M. Niemeijer, B. van Ginneken, J. Staal, M. S. A. S. Schulten, and M. D. Abramoff, "Automatic detection of red lesions in digital color fundus photographs," *IEEE Trans. Med. Image.,* vol. 24, no. 5, pp.584–592, May 2005.

[17] S. Abdelazeem, "Micro aneurysm detection using vessels removal and circular Hough transform", *In IEEE Proc. 19th Natl. Radio Sci. Conf.,* Alexandria, Egypt, Mar. 2002, vol. 1-2, pp. 421–426.