An Efficient Automated System for Detection of Diabetic Retinopathy from Fundus Images Using Support Vector Machine and Bayesian Classifiers

K.Narasimhan
Assistant Professor
Department of Electronics and Communication Engineering.
School of Electrical and Electronic Engineering.
SASTRA University.
Thanjavur,TamilNadu,India.
knr@ece.sastra.edu

V.C.Neha
Department of Electronics and Communication Engineering.
School of Electrical and Electronic Engineering.
SASTRA University.
Thanjavur,TamilNadu,India.
vcneha1989@gmail.com

K.Vijayarekha
Associate Dean-EEE
School of Electrical and Electronic Engineering.
SASTRA University.
Thanjavur,TamilNadu,India.
vijayarekha@eee.sastra.edu

Abstract - **The preliminary signs of diabetic retinopathy include micro aneurysms, haemorrhages and exudates. Early diagnosis and timely treatment can prevent vision loss in patients with long term diabetes. In this paper we used two algorithm based on filtering operations, morphological transformation and region growing method to extract features for detection of micro aneurysms, haemorrhage and non linear diffusion segmentation followed by colour histogram based clustering techniques is used to differentiate hard and soft exudates. Experimental evaluation of the algorithm has been done with images collected from Deepam Eye Hospital, Chennai, Tamilnadu, India and a database consisting of 77 abnormal and 20 normal images was created. In addition performance of the proposed algorithm is also verified on the publically available DIARETDB0 database. Based on the features obtained, each image is classified as normal or abnormal with Support Vector Machine, Bayesian Network. Classification rate of 95% is obtained with SVM and 90% with Bayesian classifier.**

Key words used **- Diabetic retinopathy, Micro aneurysm, Exudates, Haemorrhage, Support Vector Machine, Bayesian Classifier.**

I. INTRODUCTION

Diabetic retinopathy is retinopathy (damage of retina) in patients with long term diabetes, which leads to blindness. Diabetic retinopathy affects up to 80% of all patients who have diabetes for 10 years or more. 90% of these new cases could be reduced if there is proper treatment and monitoring of the eyes. There are two types of diabetic retinopathy namely proliferative and non proliferative diabetic retinopathy.

Non proliferative diabetic retinopathy is the first stage of diabetic retinopathy. Micro aneurysms appear in this stage.

The blood vessels become blocked and leads to retinal haemorrhages. As the disease progresses, severe non proliferative diabetic retinopathy enters an advanced, or proliferate. New blood vessels grow along the retina and vitreous humour that fills the inside of the eye. Later on these vessels bleed and destroy the retina. The ophthalmoscopic image of diabetic retinopathy contains red dots (micro aneurysms), haemorrhages, exudates and vascular proliferations.

Exudates are of two types namely hard exudates and soft exudates. Hard exudates are small, yellow or white waxy glistening patches with discrete margins. When hard exudates encroach on the macula vision is affected. In the case of severe hypertensive retinopathy cotton wool exudates or soft exudates are present.

Micro aneurysms are saccular pouches caused due to the distension of capillary walls. They are the first clinical sign of retinopathy.

Haemorrhages occur in the compact middle layer of the retina and appear as blots. Haemorrhages occur as flame shaped structures in the superficial nerve fiber layer and are related to severe hypertension.

II. RELATED WORK

A number of methods have been proposed for the detection of micro aneurysms. Lay et al [1] presented detection of micro aneurysms from angiograms using top hat transformation which eliminates the vasculature present leaving possible micro aneurysm. Matched filtering approach for the detection of micro aneurysms was presented by Spencer et al [2]. Deepak Vallabha et al [3] in their work proposed a method for detection and classification of vascular abnormalities in diabetic retinopathy. The vascular abnormalities are detected using scale and orientation selective Gabor filter banks. The input image is filtered through Gabor filter bank consisting of several filters, filtering operation is Fourier domain. The method classifies

978-1-4673-0210-4/12/$31.00 ©2012 IEEE

the retinal images as mild or severe based on output obtained from Gabor filter.

Another pattern classification using decision rules based on Bayes optimality criterion was put forward for the detection of micro aneurysms by Palavi Kahai et al [4]. The hypotheses supported are the null hypothesis (specifies presence of micro aneurysms) and alternate hypothesis (specifies absence of micro aneurysms). Each hypothesis has a probability density function associated with it. Each case represents a Gaussian distribution the feature value tends to a particular value but with little variation.

Keerthi Ram et al [5] devised a method of indirect detection of micro aneurysms .Instead of detection of micro aneurysms which may lead to false positives, non micro aneurysm structures are detected and rejected in two rejection stages. The first rejection stage uses nearest mean classifier, the second stage uses hyper cuboid classifier. The micro aneurysms that remain are given a numeric similarity score based on similarity to true micro aneurysms.

Alireza Osareh et al [6] proposed a method for detection of exudates in retinal images using fuzzy c-means clustering following colour normalisation and colour enhancement. The segmented images form a dataset of regions. To classify into exudates and non exudates features are extracted and a genetic based algorithm is used to rank features and identify features that gives best classification results. The feature vectors are classified using multilayer neural network.

Akara Sopharak et al [7] developed a method for exudates detection from retinal images using fuzzy C-means clustering. Contrast enhancement pre processing is applied before four features are extracted namely intensity, standard deviation on intensity, hue and a number of clustering methods. The first result was fine tuned with morphological techniques and detection results were validated by comparing hand drawn ground truths of expert ophthalmologists.

Meindert Niemeijer et al [8] discussed a method to detect exudates, drusen and cotton wool spots which employed a machine learning-based automated system. Retinal images of three hundred patients diagnosed with diabetes were selected from diabetic retinopathy tele diagnosis database. A computer program based on machine learning was developed to differentiate among drusen, exudates and cotton- wool spots. The performance of the automated system was evaluated using the false positive rate and the true positive rate.

Akara Sopharak et al [9] devised a method for detection of exudates using mathematical morphology methods. Pre processing involves converting the image to HSI colour space. A median filtering operation was followed by contrast adaptive equalisation. Followed by optic disc elimination and exudates detection using local variation operator this is followed by thresholding and reconstruction using marker image.

Akara Sopharak et al [10] devised a machine learning based approach for automatic exudates detection using naive Bayes and support vector machine classifier. The naive Bayes model is trained using 15 features. To obtain the best SVM, they started with the best feature set from the naive Bayes classifier and appended the removed features to the classifier. For each combination of features they carried out a grid search to find the best combination of hyper parameters (tolerance for training errors and radial basis function width). They compared the best naive Bayes and SVM classifiers to a baseline nearest neighbour classifier employing the best feature from both classifiers. They proved that naive Bayes and SVM classifiers executed better than NN classifier.

C.Jayakumari et al [11] devised method for exudates detection using contextual clustering and fuzzy art neural network. Here the retinal images are first subjected to pre processing for colour normalisation and contrast enhancement. Contextual clustering algorithm is used to segment retinal image. The optic disc is located and eliminated. The detected candidates s are classified as exudates or non exudates based on convex area, solidity and orientation. The modular neural network is trained using 25 images.

III. MATERIALS AND METHODS

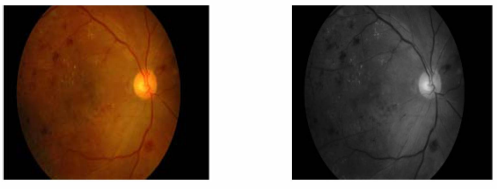
A. Materials

A total of colour fundus images with resolution of 96x96 dpi and 768 X 576 pixels were obtained from Deepam eye hospital, Chennai. The retinal photographs were taken using Topcon fundus camera model TRC50X. From the images taken 20 were identified as normal and 77 as abnormal. A database was formed, which has information of the type and location of abnormalities with the guidance of an ophthalmologist. This database serves as the ground truth for the classification process.

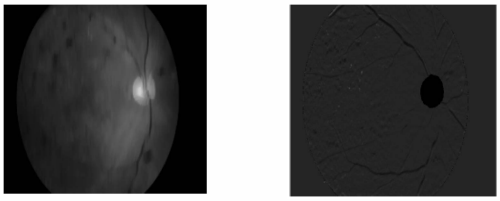
IV. DETECTION OF MICRO ANEURYSMS AND HAEMORRHAGES

A. Pre processing

To reduce the effect of intensity variation in the background, green plane of the image is extracted. The green plane image is median filtered and is then normalised by dividing the green plane image by the median filtered image [12]. The compensated normalised image is then input to the next stage.



(a) (b)



(c) (d)

Figure 1(a) Fundus image. (b) Green plane image

(c) Median filtered image (d) Compensated image.

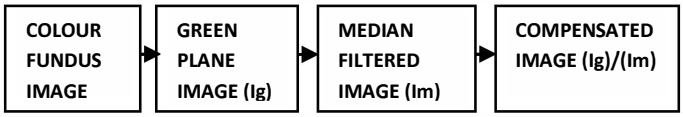


Figure 2 .Pre processing

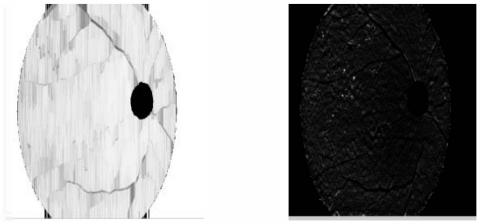
B. Morphological filtering

This is done to isolate dark circular objects in the green plane image. The geometric properties are used as the criteria for isolating the micro aneurysms and blood vessels that are same in colour. The micro aneurysms are circular whereas the blood vessels are linear in nature.

A morphological white top hat transformation is used to enhance and isolate the micro aneurysm. The white top hat transform is used to isolate structures smaller than the structuring element used. It is mainly used to isolate structures that are brighter than the back ground. The top-hat transformed image is got by obtaining the difference between the input image and opened image.

Ith=I-Iopen

The image is opened using a linear structuring element as the micro aneurysms occur in all directions. The structuring element is rotated in twelve different orientations in order to efficiently extract the small circular structures from the image. And as mentioned earlier the size of the structuring element is chosen to be larger than the largest circular feature to be segmented. Morphologically opening the image removes small objects from the foreground. As a result subtracting the maximum of the opened image from the input pre processed image isolates the circular structures from the image.



(a) (b)

Figure 3 (a) Opened image (b) Top-hat transformed image

C. Morphological Reconstruction

The morphological top-hat transformation is then succeeded by the geodesic reconstruction by dilation [13]. The top hat transformation may exclude small curved vessels and consequently the top hat transformed image contains curved vessels in addition to micro aneurysms.

Figure 4.Morphological filtering and reconstruction.

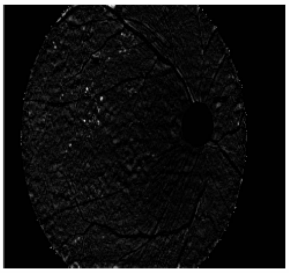


Figure 5. Reconstructed image.

Therefore a geodesic reconstruction is carried out to recover these linear features. Geodesic dilation involves two images a marker and mask image. Here the marker image is got by dilating the opened image using an isotropic structuring element. A square structuring element is used. Then geodesic reconstruction is carried out using the top-hat transformed image as the mask and the dilated image as the marker. The image obtained is enhanced using histogram equalisation.

D. Thresholding

Micro aneurysms are detected from the reconstructed image employing thresholding technique. A conventional adaptive technique is used to set the threshold. The threshold is set to a low value but this result in false detection. Therefore the threshold value is optimised using the thresholding formula.

$$Threshold = b\ m_4 \qquad (1)$$

Where m_4 is the fourth order moment.

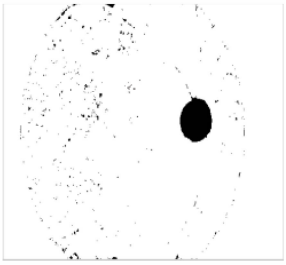


Figure 6.Thresholded image.

E. Region Growing

Region growing is region based image segmentation to partition image into regions that exploits the fact that pixels which are close together have similar gray values. The region growing process is done by first selecting the seed pixel. The

neighbouring pixels are checked for similarity to the seed, this is repeated for each newly added pixel. The homogeneity predicate can be based on any characteristic of the image. In this case the seeds are selected by looking over the small connected region and retaining pixel with lowest grey level. The connected region formation is accomplished using the condition [14],

$$P \leq Pseed + x(Pbg - Pseed) \quad (2)$$

P is the grey level of the pixel under test, Pseed is the gray level of the seed pixel, x is a region growing parameter value is approximately unity and Pbg is the grey level of the background pixel. The region growing process is terminated when this condition is not satisfied and a new seed pixel is chosen and the process continues.

F. Feature Extraction

The output from the previous region growing stage is analyzed to locate the micro aneurysms based on external shape and colour features.

The classification based on external shape takes into account three features namely circularity, aspect ratio and area of the target lesion. The circularity of the target should be ideally unity. Whereas the aspect ratio which is the height to width ratio should be unity and the area should contain no more than two pixels. For area containing more than two pixels can be identified as haemorrhages.

The colour feature on the other hand is the hue. And micro aneurysms are detected by the redness (low value of hue). In order to avoid false detection the detection threshold is adaptively adjusted between images using an exponential constant false alarm rate thresholding technique.

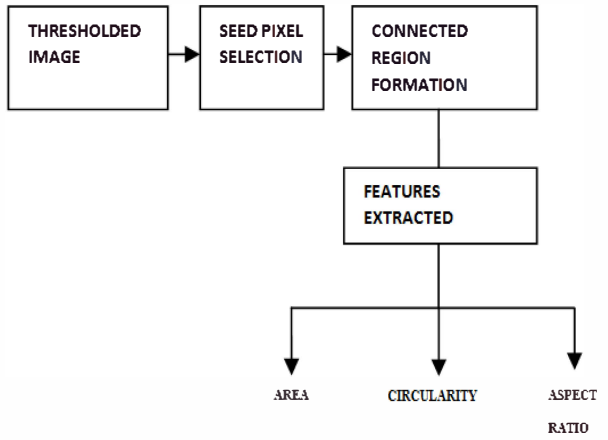
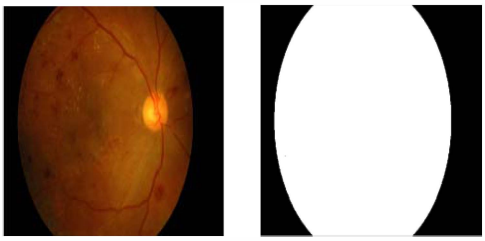


Figure 6.Feature extraction

V. DETECTION OF EXUDATES

A. Localisation of optic disk

A binarization process is carried out to detect the fundus region [15]. This is done by a thresholding method which replaces pixels with high intensity into white regions and pixels with low intensity into black regions. The properties of the connected regions in the image are measured using 'regionprop' function. The binary image is usually a logical array. Colour histogram is then applied to each individual region. The optic disc is identified as the region with maximum pixel value.



(a) (b)

Figure 7 (a) Fundus image (b) Fundus mask.

B. Segmentation

The vital step in image segmentation is the detection of edges. Edges of an image is characterised by significant changes in the image intensity. To localise the image boundaries and regions that are piece wise constant non linear diffusion segmentation is used [16]. Non linear diffusion segmentation removes low frequency noise in the segmentation process. Non linear diffusion segmentation [17] developed by Perona and Malik is employed. This is done using the equation

$$It = div(c(\nabla I)\nabla I) \quad (3)$$

Where I is the input image to the segmentation process, It is the evolution over time and c(∇I) is the decreasing function of ∇I. Segmentation is achieved by finding the edges in the smoothened input image.

Figure 7.Detection of exudates.

C. Colour histogram clustering technique

Colour histogram thresholding is used to detect exudates [18]. This is done by splitting the colour fundus image into a number of overlapping blocks. Thresholding is done to detect exudates [19] .Accurate choice of threshold value can be used to differentiate between hard and soft exudates.

The characteristics of the histogram are observed for each non overlapping block. It is seen that the histogram of the blocks containing exudates ranges approximately between 100 to 180, whereas the histogram of the blocks not containing exudates ranges between 0 to 160.

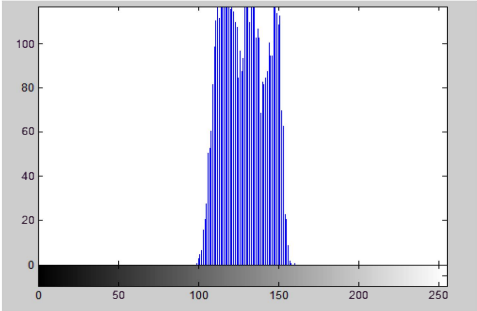


Figure 8.Histogram of block containing exudates.

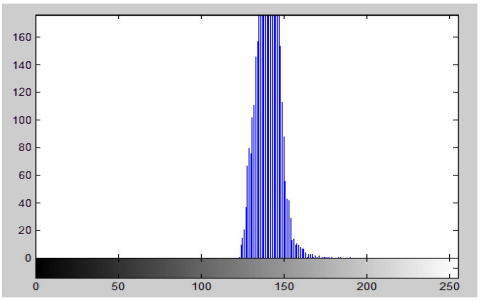


Figure 9.Histogram of block not containing exudates.

VI. FEATURE BASE CREATION

The number of testing and training set are shown in table 1. From the table, totally 50 fundus images are used to train the classifier. The proposed features area, circularity and aspect ratio and colour feature are calculated for each of the training images and for each training image range of threshold value is calculated for exudates and non exudates region and stored in an array. This array is called the Feature base and the size of the feature base is 15x2.

VII. EXPERIMENTAL RESULTS AND CLASSIFICATION

In the classification stage, pre processing operation of unknown image is done and various features are extracted to detect micro aneurysm and exudates. These features are processed with the features stored in the feature database to make the decision [20].In automated systems used for detection of objects it is essential to evaluate the performance of the algorithm. The results obtained by the system can be compared to the diagnosis made by the specialist. The test result indicate whether the disease is present or not(positive or negative).The performance of the proposed algorithm is tested in terms of True positive (TP), True negative (TN), False positive (FP), False negative (FN), sensitivity (SE), specificity (SP) and accuracy (AUC).True positive(TP) indicates that the patient suffers from the disease and test result was also positive, False positive(FP) indicates that the patient does not suffer from the disease and was diagnosed as positive, True negative(TN) indicates that the patient does not suffer from the disease and was diagnosed as negative, False negative(FN) indicates that the patient suffers from the disease but diagnosed as negative. These metrics are given as

$$SE = \frac{TP}{TP + FN} \tag{4}$$

$$SP = \frac{TN}{TN + FP} \tag{5}$$

$$AUC = \frac{TP + TN}{TP + FN + TN + FP} \tag{6}$$

The sum of the true positive and false negative rate is the number of patients suffering from the disease. And the sum of true negative and false positive is the number of patients not suffering from the disease. Sensitivity is the measure of percentage of cases detected and specificity is the measure of correctly classified healthy persons. In addition, performance of algorithm was measured with receiver operating characteristics (ROC) curves. The ROC curve plots the sensitivity against specificity of a system. ROC is an indication of the capability of a system to distinguish those images containing diabetic retinopathy and normal images. In ideal prediction we get a point in the upper left corner, or coordinate (0, 1) of the ROC space indicating 100% sensitivity and 100% specificity .Of a total of 47 images collected from the hospital we achieved a true positive value of 35, false positive value of 3 and a false negative and true negative value of 2 and 7 respectively. In the database collected from the hospital we achieved AUC of 0.95 and for DIARETDB0 0.92.

TABLE I

Category	No. of Training Set	No. of testing Set
Normal	10	10
Abnormal	40	37

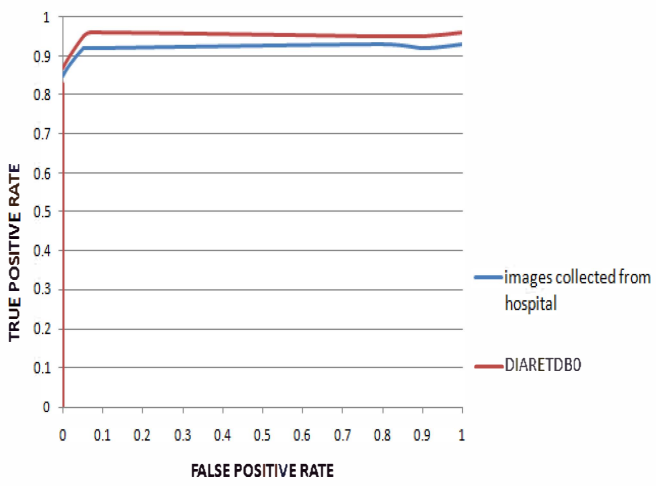


Figure 10.ROC curves: Image level performance on DIARETDB0 and images collected from hospital.

VIII. CONCLUSION

In the current work, we have presented an algorithm for the detection of micro-aneurysms and exudates. We have optimized the feature vector based on the ROC analysis and we obtained AUC of 0.95 for DIARETDB0 database and 0.92 for the database collected from hospital with SVM.SVM shows better performance in classification than Bayesian classifier. In Bayesian decision problem is posed in probabilistic term and decision made is perfect only if the all of the relevant probability values are known. In DIARETDB0 we got good results because fixed protocol is used to acquire images. One of the advantages of the proposed system is the flexibility and simplicity when compared to the other systems proposed earlier. For improvement in results obtained there are number of areas which needs further investigation. The most effective way is to use additional features and by increasing the amount of training data. The overall system performance may be further improved by using heterogeneous database and feature vector selection has to be done with different testing methods like Wilcoxon rank test, Ansari-Bradley test and the feature vector which unlikely to be useful has to be rejected.

REFERENCES

[1] B.Lay,"Analyse automatique des images angioflurographiques au cours de la reaatinopathie diabeaatique,"Ph.D thesis,Center of Mathematical Morphology,Paris School of Mines,June 1983.

[2] T.Spencer,R.P.Philips,P.Sharp and J.Forrester,"Automated detection and quantification of micro aneurysms in fluorescein angiograms,"Graefe's Archive for Clinical and Experimental Opthalmology,vol.230,pp 36-41,1991.

[3] Deepak Vallabha,Ramprasath Dorairaj,Kamesh Namuduri,Hilary Thompson,"Automated Detection and Classification of Vascular Abnormalities in Diabetic Retinopathy",IEEE Conference on Signals,Systems and Computers,Vol.2,pp.1625-1629.2005.

[4] Pallavi Kahai,Kamesh Rao Namuduri,Hilary Thompson,"Decision Support for Automated Screening of Diabetic Retinopathy",IEEE Conference on Signals, Systems and Computer.Vol.2,pp. 1630-1634.2005.

[5] Keerthi Ram,Gopal Datt Joshi and Jayanthi Sivasamy,"A Successive Clutter-Rejection based Approach for Early Detection of Diabetic Retinopathy",IEEE Transactions on Biomedical Engineering.2011.

[6] Alireza Osareh,Bita Shadgar,Richard Markham,"A Computational-Intelligence-Based Approach for detection of exudates in Diabetic retinopathy images",IEEE Transactions on Information Technology in Biomedicine,Vol.13,NO.4,July 2009.

[7] Akara Sopharak,Bunyarit Uyyanonvara,Sarah Barman,"Automatic Exudate Detection from Non-dilated Diabetic Retinopathy Retinal Images Using Fuzzy C-mean Clustering",Journal of Sensors,Vol.9,No.3,pp.2148-2161,2009.

[8] Meindert Niemeijer,Bram van Ginnekan,Stephan R.Russell,Maria S.A.Suttorp-Schulten,Michael D.Abramoff,"Automated Detection and Differentiation of Drusen,Exudates and Cotton-Wool Spots in Digital Colour fundus Photographs for Diabetic Retinopathy Diagonis",Investigative Opthalmology and Visual Science,Vol 48,pp.2260-2267,2007.

[9] Akara Sopharaka,Bunayarit Uyyanonvaraa,Sarah Barmanb,Thomas H.Williamspnc,"Automated Detection of diabetic retinopathy exudates from non-dialated retinal images using mathematical morphology methods",Computerised Medical Imaging and Graphics ,Vol.32,No.8,pp.720-727,2008.

[10] Akara Sopharak,Matthew N.Dailey ,Bunyarit Uyyanonvara,Sarah Barman,Tom Williamson,Khine Thet Nwe,Yin Aye Moe,"Machine learning approach to automatic exudates detection in retinal images from diabetic patients",Journal of Modern Optics,2009.

[11] C.Jayakumari and T.Santhanam,"Detection of Hard Exudates for Diabetic Retinopathy using Contextual Clustering and Fuzzy Art Neural Network" .Asian Journal of Information Technology,Vol.6,No.8,pp 842-846,2007.

[12] Ahmed Wasif Reza, C.Eswaran, Subha Hati,"Automatic Tracing of Optic disc and Exudates from Colour Fundus Images Using Fixed and Variable Threshold",Journal of Medical Systems,Vol.33,No. 1,pp. 73-80,February 2009.

[13] M.J. Cree, J.A. Olson, K. C.Mchardy, P.F.Sharp, and J.V. Forrester,"A fully automated comparative microaneurysm digital detection system," Eye,Vol. 11,pp. 622-628, 1997.

[14] Sinthanayothin C,Boyce JF,Cook HL,Williamson TH.,"Automated localisation of the optic disc,fovea and retinal blood vessels from digital colour fundus images",Br J Opthalmol.,Vol. 83,pp. 902-910,1999.

[15] T. Spencer, J.A. Olson, K.C. Mchardy, P.F. Sharp, J. V. Forrester, "An image-processing strategy for the segmentation and quantification of micro aneurysms in fluorescein angiograms of the ocuar fundus",Computers and biomedical research, Vol.29,pp. 284-302, 1996.

[16] L.Alvarez, P.L. Lions and J.M. Morel,"Image Selective Smoothing and Edge Detection by Nonlinear Diffusion II",SIAM Journal of Numerical Analysis,vol.29,no. 3,pp. 845-866,1992.

[17] P.Perona and J.Malik,"Scale space and edge detection using anisotropic diffusion",IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.12, No.7,pp.629-639,July 1990.

[18] Otsu .N,"A Threshold selection method from gray –level histograms.IEEE Transactions on Syst.Man. and Cybern. SMC-9,pp. 62-66.1979.

[19] O. Chapelle, P.Haffner and V. Vapnik."SVMs for histogram-based image classification. IEEE Transactions on Neural Networks,vol.10,pp. 1055-1065,1999

[20] Osareh A,M. Mirmehdi, B.T.Thomas and R. Markham, 2002. Classificatiion and Localisation of Diabetic Related Eye Disease. Proc. 7th Eur. Conf. Comput. Vision, pp. 502-516.