International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 3, Issue 5, May 2016

Automated Detection and Classification of Diabetic Retinopathy using Morphological processing and Support Vector Machine

[1] Harini R [2] Sheela N
[1] Student (M. Tech), [2] Assistant Professor,
[1] [2] Department of Information Technology,
Sri Jayachamarajendra College of Engineering [SJCE], Mysuru
[1] harini.r42@gmail.com [2] sheelaraon@gmail.com

Abstract: **The Digital image processing helps the ophthalmologists in distinguishing the vascular irregularities in order to detect some disorders related to retina like Diabetic Retinopathy (DR), Age-related Macular Degeneration (AMD), and Glaucoma etc., which can cause visual impairments. The retinal fundus images of the patients are procured by capturing the fundus of the eye with a digital fondues camera. The manual method of observing several retinal fundus images by the ophthalmologists is time consuming. Therefore, a computer assisted automated method is very helpful. In this paper a method for DR detection by utilizing Fuzzy C-Means (FCM) clustering and morphological image processing is proposed. The image pre-processing includes image resizing, CLAHE, contrast adjustment, gray and green channel extraction from the color fundus image. The classification by Support Vector Machine (SVM) classifier using selected features achieves an Accuracy of 87.50%, Sensitivity of 83.33%, and Specificity of 90%.**

Keywords: Diabetic Retinopathy (DR), Blood vessel, Micro aneurysms, Exudates, Support Vector Machine (SVM).

I. INTRODUCTION

The disease of interest here, the Diabetic Retinopathy (DR) has to be detected to avoid the vision loss of the patient. It occurs when there is an increased blood glucose level damagingthe capillaries causing blood leakage into the retina. Thisleakage is seen in retinal images as the abnormal features such as microaneurysms, hemorrhages, hard exudates, cotton wool spots etc. Depending on the presence of these abnormal features, the stages of DR can be identified. The retinal image manual screening is time consuming and less accurate due to human error. Therefore, there isa need for Automated DR detection systems which can provide simple approachfor classifying the images as normal or DR with increased accuracy and efficiency.

II. LITERATURE REVIEW

The survey done for achieving this work is explained as following. Shilpa Joshi,et.al. [1] have extracted blood vessels using morphological operations. Kittipol Wisaeng, et.al. [2] have presented a new method for detecting the exudates pathologies of DR using Fuzzy C-Means (FCM) clustering and morphological methods. M. Niemeijer, J. Staal, et.al. [3] have compared vessel segmentation algorithms. Priya.R, et.al. [4] have investigated and proposed a computer based system to identify normal, Non-proliferative diabetic retinopathy (NPDR) and Proliferative retinopathy (PDR) using PNN (Probabilistic Neural Network), Bayes theory and SVM. Madhura Jagannath et.al. [5] have presented an automated system for the detection of Diabetic Retinopathy by extracting blood vessels and exudates using morphological operations and texture features using GLCM. Oliver Faust, et.al. [6] have reviewed on algorithms for classifying DR as normal, NPDR and PDR using ANN (Artificial Neural Network) and SVM.

M. Ponni Bala, et.al. [7] have proposed an automated technique for detecting and classifying the retinal images into exudates and non-exudates with severity of disease using an automated Fuzzy Inference System (FIS).Archana G, et.al. [8] have presented an automated method for the detection of bright lesions (exudates) and classified as normal or abnormal based on the features obtained from the image.

III. PROPOSED METHOD

The objective of the proposed method is to develop an automated system for Diabetic Retinopathy detection and classification. The Fuzzy C-Means

Clustering and Morphological operations are used to extract Blood vessel, Microaneurysms and Exudates. The classification is done depending on the area of blood vessel, area of microaneurysms, and area of exudates using SVM classifier. The input image is first resized to 576x720, and then CLAHE, contrast enhancement is applied in pre-processing stage to improve the quality of the image. The general flowchart of the proposed method is as shown in figure 1 below.

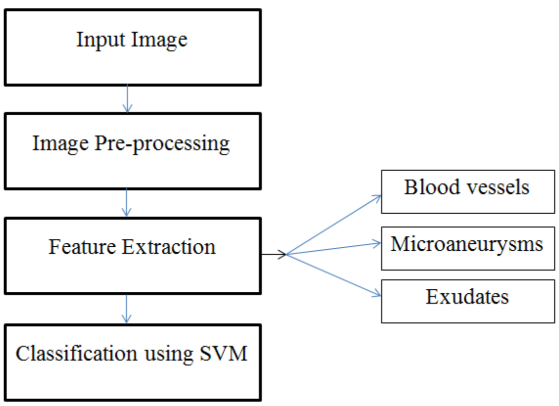

Fig. 1 Flowchart of proposed method

Blood Vessels segmentation**:** The segmentation of Blood vessels from the fundus eye images is performed by using the Fuzzy C-Means (FCM) clustering and the fine tuning is performed by using Morphological operations.

The FCM clustering is an iterative overlapping algorithm where Euclidean distance between the feature vectors and the cluster centers are used to define similarity. First step cluster number (C), number of iterations (K), and epsilon (e) the value for stop condition are initialized.

The objective function J is given by theequation 1.

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} {U_{ij}}^2 \ \|X_i - C_j\|^2 \quad (1)$$

Now, the fuzzy partition matrix U_{ij}, U = (u_{ij}) (U (0)) is initialized by generating random numbers in the range 0-1. The cluster center C_j is calculated using equation 3. The equation for U_{ij} is given by equation 2. Where $U_{ij}$is the degree membership of X_i, in the cluster j, and X_i is the i^{th} measured data. The update of membership U_{ij} and the cluster centers C_j are carried out by using the equation 2 and 3 respectively.

$$U_{ij} = \frac{1}{\sum_{K=1}^{C} \left(\frac{\|X_i - C_j\|}{\|X_j - C_k\|} \right)^2} \quad (2)$$

$$C_j = \frac{\sum_{i=1}^{N} {U_{ij}}^2 \ X_i}{\sum_{i=1}^{N} {U_{ij}}^2} \quad (3)$$

The objective function, J is updated as and when U_{ij} and C_j are updated until the difference between adjacent values of the objective function is less than the epsilon. When this condition is met, the iteration will stop. The condition is mathematically represented as in equation 4.

$$max_{ij} = \{|{U_{ij}}^{K+1} - {U_{ij}}^{K}|\} \ < \ \varepsilon \quad (4)$$

The result of blood vessels extraction after fine tuning the FCM clustering result is as shown in figure 2 below. The area of the blood vessels is calculated from this result.

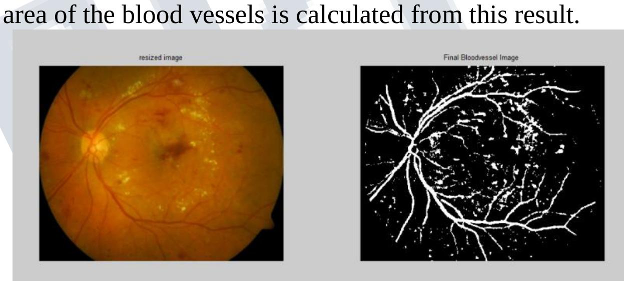

Fig. 2 Result of Blood vessels segmentation

Microaneurysms detection: The process involves removal of blood vessels area and larger area in order to extract microaneurysms area. The morphological operations are applied; optic disk (OD) is removed by creating and applying optic disk mask. The circular and rectangular borders are formed to remove border noise. By using logical AND operation for image with noise and the CLAHE output image, the bright regions (exudates) are eliminated. The blood vessels are eliminated by performing logical AND operation between the previous result and the result of removing small noise areas from inverted CLAHE applied image. The final result of microaneurysms extraction is as shown in figure 3 below. The area of microaneurysms is calculated from this result.

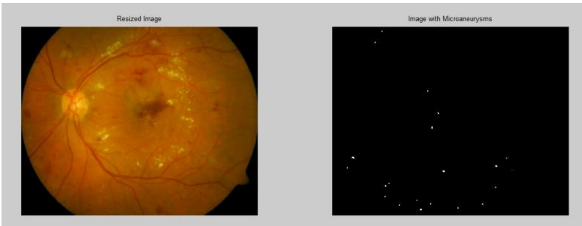

Fig. 3 Result of Microaneurysms detection

Exudates detection: The main problem here is, the optic disk and exudates have same intensity. Therefore, the OD area is eliminated. Next, the blood vessels are removed by using the morphological operations. The circular border for eliminating border noise is formed by using Canny edge detection and applying morphological operations. To remove the OD, first a meshgrid of image size is formed. After finding the brightest point in the image, its co-ordinates are marked as center and a circle is drawn using formula of circle as in equation 5.

$$R^2=(x-h)^2+(y-k)^2 \qquad (5)$$

The Circular border is as shown in figure 4 and OD removal mask is as shown in figure 5.

Fig. 4 Circular border

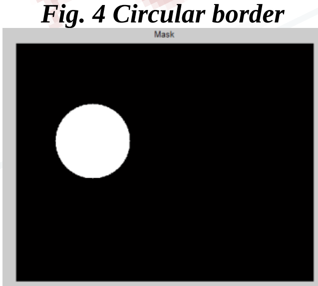

Fig. 5 OD mask

The exudates detection is done by adding the results of two morphological operation based techniques. One method uses Top-Hat transformation where background exclusion is performed for illumination correction and other method uses blood vessels removal, segmentation and logical AND operation. The final result of the detection is as shown in the figure 6 below. The area of exudates is calculated from this result.

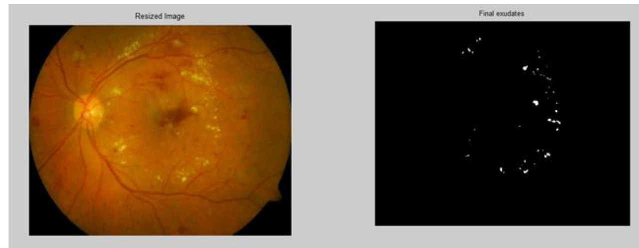

Fig. 6 Result of Exudates detection

Therefore, depending on the area of blood vessels, area of Microaneurysms and area of Exudates the image is classified as normal or DR.

IV. CLASSIFICATION USING SVM

The classification of the images as normal or DR is performed by using the Support Vector Machine (SVM) classifier, a supervised learning algorithm. The SVM system uses hypothesis space of linear functions. In high dimension feature space the kernel functions are used for learning. The data is trainedwith a learning algorithm from optimization theory of Lagrange. The bias for learning is derived from Generalization i.e. statistical learning theory. The decision boundary is nothing but the hyper plane which separates the positive and negative data points and this separation must be maximum. A linear classifier has the form as given below in equation 6.

$$f(x) = w` x + b \qquad (6)$$

Where w is the weight vector and has to be normal to the median of the margin separation (to hyper plane); b is the bias.

We have to add additional constraints to calculate b and w` to make decision.

The optimization problem (used to find the optimal hyper plane) and the decision function can be expressed in dual form which depends only on dot products between vectors. The dual representation of the decision function is as in equation 7.

$$f(x_i)= \mathrm{sign}(\Sigma\, y_i\, \alpha_i\, (w`.\, x_i` + b)) \qquad (7)$$

Given training data (x_i, y_i) for i = 1. . . N, with xi ∈ R^d and $y_i \in \{-1, 1\}$, learn a classifier f(x) such that

$$f(x_i) \quad \geq 0 \; y_i =+1$$

$< 0 \ y_i = -1$

i.e. $y_i f(x_i) > 0$ for a correct classification.

The margin of separation is given by equation 8.

$(w / \|w\|).(x^+ - x^-) = w`(x^+ - x^-) / \|w\| = 2 / \|w\|$ (8)

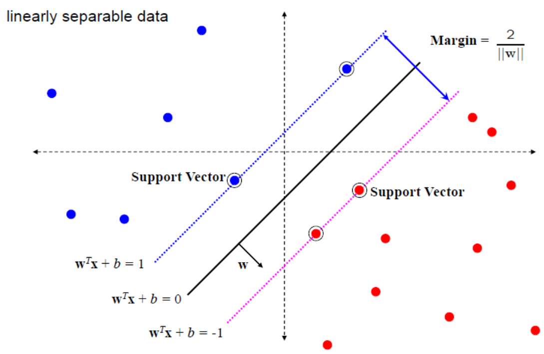

Fig. 7 Illustration of SVM for linearly separable data points

The points nearest to the separating hyper plane are called Support Vectors as shown in figure 7. Thex is support vectors. Only they determine the position of the hyper plane. All other points have no influence. Mathematically the weighted sum of the Support Vectors is the normal vector of the hyper plane.

Learning the SVM can be formulated as an optimization. So we have,

Max $(2 / \|w\|)$ subject to $w` x_i + b \geq 1$ if $y_i = +1$

≤ -1 if $y_i = -1$ for $i = 1. . . N$

Or equivalently,

Min $(\|w\|2)$ subject to yi $(w` x_i + b) \geq 1$ for $i = 1. . . N$

Now we have a quadratic optimization problem and we need to solve for w and b. To solve this we need to optimize the quadratic function with linear constraints. The solution involves constructing a dual problem and where a Lagrange's multiplier αi is associated. We need to find w and b such that Φ (w) $=½ |w'\|w|$ is minimized; And for all $\{(x_i, y_i)\}$: $y_i (w * x_i + b) \geq 1$. Thus the maximal margin can be found by maximizing the Lagrangian. The Lagrange function is given by equation 9.

$L= ½ \|w`\|^2 - \Sigma \alpha_i [y_i (w`. x_i` + b)-1]$ (9)

Differentiate L partially with respect to w`

The result is $w`= \Sigma \alpha_i y_i x_i`$ (10)

i.e., the decision vector w is linear sum of samples.

Next differentiate L with respect to b

The result is $\Sigma \alpha_i y_i = 0$ (11)

Substituting equation 10 and 11 back in equation 9, we get result as in equation 12.

$L= \Sigma \alpha_i - (1/2) (\Sigma \Sigma \ \alpha_i \alpha_j y_i y_j (x_i`. x_j`))$ (12)

The optimization depends only on dot product $(x_i`. x_j`)$. These give the sample vectors.

The decision rule will look on this with unknown point u.

If $\Sigma \alpha_i y_i x_i`. u` + b \geq 0$ then the sample is positive else negative.

To minimize L, $\alpha_i (y_i (w`. x_i` + b)-1) = 0$

Thus, only support vectors haveαi not equal to zero. Therefore, w' can be calculated from the equation and b can be found from the equation 13.

$y_i (w`. x_i` + b)-1 = 0$ (13)

Now for the linearly non-separable samples, to allow for linear separation kernels are used to non-linearly map the input data to a high-dimensional space. The new mapping is then linearly separable.A kernel function is defined as a function that corresponds to a dot product of two feature vectors in some expanded feature space. This mapping is defined by the Kernel as in equation 14.

$K(x,y) = \Phi(x) \cdot \Phi(y)$ (14)

V. RESULTS AND DISCUSSION

The Blood vessel segmentation is carried out by FCM clustering and Morphological operations. The Microaneurysms and Exudates are detected using Morphological processing. The classification is carried outusing SVM classifier. The image data set taken for classification is total 40 in numbers. A total number of 24 images are taken into training data set. Among which 9 are normal images and 15 are abnormal images. This training data set is trained using SVM classifier for the features taken to be as Blood vessel area, Microaneurysms area, Exudates area. For the testing, 16 images are considered in testing data set; of which 6 are normal and 10 are abnormal images. These images in testing data set are tested using SVM classifier. As aresult of classification, there was correct classification of 14 test images, and 2 images were wrongly classified. The performance of the classifier is evaluated and it achieves an Accuracy of 87.50%, Sensitivity of 83.33% and Specificity of 90% in classifying the test images.

The performance of the classifier can be improved by including more number of features for classification.

VI. CONCLUSION

The proposed method where Fuzzy C-means clustering and Morphological processing are used for segmentation of Blood vessels, and Morphological operations to detectMicroaneurysms and Exudates from the retinal fundus image is applied on images from DIARETDB0 and DIARETDB1 databases. The images are classified by using SVM classifier. The Blood vessel area, Microaneurysms area, Exudates area are the features taken for classification. The proposed method achieves an Accuracy of 87.50%, Sensitivity of 83.33% and Specificity of 90%.

REFERENCES

[1] Shilpa Joshi and Dr P.T. Karule. "Retinal Blood Vessel Segmentation", International Journal of Engineering and Innovative Technology (IJEIT), Volume 1, Issue 3, March 2012.

[2] Kittipol Wisaeng, Nualsawat Hiransakolwong and EkkaratPothiruk. "Automatic Detection of Exudates in Diabetic Retinopathy Images", Journal of Computer Science 8 (8): 1304-1313, 2012 ISSN 1549-3636, Science Publications © 2012.

[3] M. Niemeijer, J. Staal, B. v. Ginneken, M. Loog, and M. D. Abramoff, J. Fitzpatrick and M. Sonka, Eds., "Comparative study of Retinal Vessel Segmentation methods on a new publicly available database," in SPIE Med. Imag., vol. 5370, pp. 648–656, 2004.

[4] Priya.R, Aruna.P, "Diagnosis of Diabetic Retinopathy using Machine Learning Techniques", ICTACT JOURNAL ON SOFT COMPUTING, Volume: 03, Issue: 04, July 2013.

[5] Madhura Jagannath Paranjpe, Prof. M N Kakatkar, "Automated Diabetic Retinopathy Severity Classification using Support Vector Machine", International Journal for Research in Science & Advanced Technologies, Volume-3, Issue-3, 086-091, ISSN 2319-2690, May-June 2013.

[6] Oliver Faust & Rajendra Acharya U. & E. Y. K. Ng & Kwan-Hoong Ng & Jasjit S. Suri, "Algorithms for the Automated Detection of Diabetic Retinopathy Using Digital Fundus Images: A Review", Springer Science+Business Media, J Med Syst, 6 April 2010.

[7] M. Ponni Bala, S. Vijayachitra, "Computerised Retinal Image Analysis to Detect and Quantify Exudates Associated with Diabetic Retinopathy", International Journal of Computer Applications (0975 – 8887), Volume 54– No.2, September 2012.

[8] Archana .G, et.al., "Abnormality Detection and Its Severity Classification in Retinal Images", International Journal of Research in Engineering & Advanced Technology (IJREAT), Volume 1, Issue 1, ISSN: 2320 – 8791, March, 2013.