An Efficient Approach for   Detection of Exudates in Diabetic Retinopathy Images Using Clustering Algorithm

G.S.Annie Grace Vimala[1], Dr.S.Kaja Mohideen[2]

[1](ECE, St.Joseph's Institute of Technology, India)
[2](ECE, B.S.Abdur Rahman University, India)

Abstract : *Diabetic retinopathy is a kind of disorder which occurs due to high blood sugar level. This disorder affects retina in many ways. Blood vessels in the retina gets altered . Exudates are secreted, hemorrhages occur, swellings appear in the retina. Diabetic Retinopathy (DR) is the major cause of blindness. Automatic Recognition of DR lesions like Exudates, in digital fundus images can contribute to the diagnosis and screening of this disease. In this approach, an automatic and efficient method to detect the exudates is proposed. The real time retinal images are obtained from a nearby hospital. The retinal images are pre-processed via. Contrast Limited Adaptive Histogram Equalization (CLAHE). The preprocessed colour retinal images are segmented using K-Means Clustering technique. The segmented images establish a dataset of regions. To classify these segmented regions into Exudates and Non-Exudates, a set of features based on colour and texture are extracted. Classification is done using support Vector Machine This method appears promising as it can detect the very small areas of exudates.*

Keywords – *Diabetic Retinopathy, Exudates, fundus image, k means clustering, SVM*

I.   INTRODUCTION

The retina is the inner and most important layer of the eye. It is composed of several important anatomical structures which can indicate various diseases. Cardiovascular disease such as stroke and myocardial infarction can be identified from retinal blood vessels. Diabetic Retinopathy is the common retinal complication associated with diabetes. It is a major cause of blindness in both middle and advanced age groups. The International Diabetes Federation reports that over 50 million people in India have this disease and it is growing rapidly (IDF 2009a) [2]. The estimated prevalence of diabetes for all age groups worldwide was 2.8% in 2000 and 4.4% in 2030 meaning that the total number of diabetes patients is forecasted to rise from 171 million in 2000 to 366 million in 2030 [3]. Therefore regular screening is the most efficient way of reducing the vision loss.

Diabetic Retinopathy is mainly caused by the changes in the blood vessels of the retina due to increased blood glucose level. Exudates are one of the primary sign of Diabetic Retinopathy [5]. Exudates are yellow-white lesions with relatively distinct margins. Exudates are lipids and proteins that deposits and leaks from the damaged blood vessels within the retina. Detection of Exudates by ophthalmologists is a laborious process as they have to spend a great deal of time in manual analysis and diagnosis. Moreover, manual detection requires using chemical dilation material which takes time and has negative side effects on patients. Hence automatic screening techniques for exudates are preferred.

1.1 *Overview of State of Art*

Alireza Osareh *et al* [2] proposed a method for automatic identification of exudates based on computational Intelligence technique The colour retinal images were segmented using fuzzy c-means clustering. Feature vector were extracted and classified using multilayer neural network classifier. Akara Sopharak *et al* [6] reported the result of an automated detection of exudates from low contrast digital images of retinopathy patients with non-dilated pupils by Fuzzy C-Means clustering. Four features such as intensity, standard deviation on intensity, hue and a number of edge pixels were extracted and applied as input to coarse segmentation using FCM clustering method..

Niemeijer *et al* [7] distinguished the bright lesion like exudates, cotton wool spots and drusen from colour retinal images. In the first step, pixels were classified, resulting in a probability map that included the probability of each pixel to be part of a bright lesion.

Walter *et al* [4] identified exudates from green channel of the retinal images according to their gray level variation. The exudates contours were determined using mathematical morphology techniques. However the author ignored some types of errors on the border of the segmented exudates in their reported performances and did not discriminate exudates from cotton wool spots.

II.   Imaging And Image Acquisition

To evaluate the performance of this method, the digital retinal images were acquired using VISUCAM from CARL ZEISS MEDITEC at a nearby hospital. These retinal images are acquired through high sensitive color fundus camera with the illumination, resolution, field of view, magnification and dilation procedures kept constant.
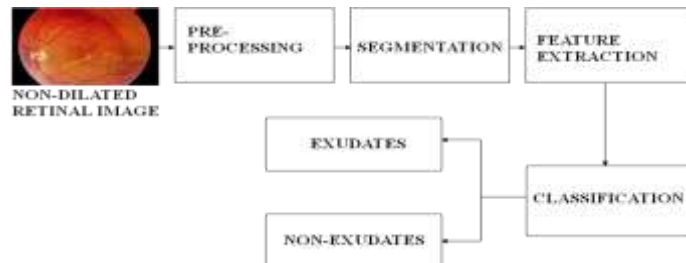

Figure 2.1 Block Diagram

2.1 Image Pre-Processing And Processing

Colour fundus images often show important lighting variation, poor contrast and noise. In order to reduce these imperfections [11] and generate images more suitable for extracting the pixel features in the classification process, a preprocessing comprising the following step is applied. 1) RGB to HSI conversion 2) Median Filtering 3) Contrast Limited Adaptive Histogram Equalization (CLAHE).

RGB to HSI Conversion:

The input retinal images in RGB Colour space are converted to HSI colour space. The noise in the images are due to the uneven distribution of the intensity (I) component.

Median Filtering:

In order to uniformly distribute the intensity throughout the image, the I-component of HSI colour space is extracted and filtered out through a 3X3 median filter.

Contrast Limited Adaptive Histogram Equalization (CLAHE):

The contrast limited adaptive histogram equalization is applied on the filtered I-component of the image [12]. The histogram equalized I component is combined with HS component and transformed back to the original RGB colour space.
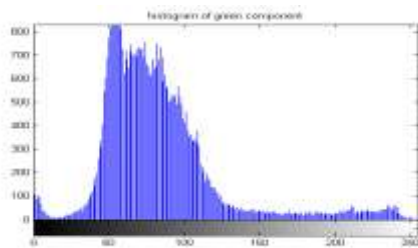

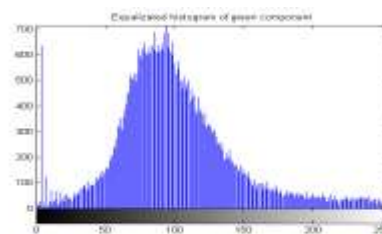Figure 2.2   Histogram of Green Component   Figure 2.3   CLAHE - Green Component

2.2 Image Segmentation Based On K-Means:

Image segmentation is the process of partitioning of an image into meaningful regions based on homogeneity or heterogeneity criteria. Image segmentation can be pixel oriented, contour oriented, region oriented, model oriented, color oriented and hybrid. In this approach, we present a novel image segmentation based on colour features from the images. The work is divided into two stages: First, enhancing the colour separation is done by extracting the a*b* components from the L*a*b* colour space of the preprocessed image. Then, the regions are grouped into a set of five clusters using K-means Clustering algorithm. By this two step process, we reduce the computational cost avoiding feature calculation for every pixel in the image[13].

The entire process can be summarized in following steps:

Step 1: Read the image. Fig 2.4 shows the example input retinal image with exudates.

Step 2: Convert the image from RGB colour space to L*a*b* colour space (Figure 2.5). L*a*b* colour space helps us to classify the color differences. It is derived from the CIE XYZ tristimulus values. L*a*b* colour space consists of a Luminosity layer L*, chromaticity layer a* indicating where the colour falls along the red-green axis, chromaticity layer b* indicating where the colour falls along the blue-yellow axis. All of the colour information is in the a* and b* layer. The difference between two colours can be measured using the Euclidean distance.
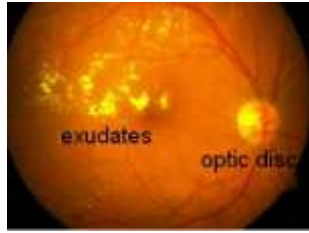
Figure 2.4 Input Color retinal image.   Figure2.5 CIE L*a*b colour space conversion

Step 3: Segment the colours in a*b* space using K-means clustering. Clustering is a way to separate groups of objects. K-Means Clustering treats each object as having a location in space. It finds partition such that objects within each cluster are close to each other as possible and as far from other objects in other clusters as possible. The algorithm requires that we specify the number of clusters to be partitioned and a distance metric to quantify how close two objects are to each others. Since the colour information exist in the a*b* space, our objects are pixels with a* and b* values. Use K-means to cluster the objects into five clusters using the Euclidean distance metric.

Step 4: Label every pixel in the image using the result from K-means .For every objects in the input, K-means returns an index corresponding to a cluster. Label every pixel in the image with its cluster index.

Step 5: Create images that segment the images by colour.

Step 6: Since the Optic Disc and Exudates are homogenous in their colour property, cluster possessing Optic Disc is localized for further processing.

III.   Feature Extraction:

To classify the localized segmented image into exudates and Non-exudates, a number of features based on colour and texture are extracted using Gray Level Co-occurrence Matrix (GLCM) . GLCM is a tabulation of how often different combination of pixel brightness values occur in a pixel pair in an image. Each element (i, j) in GLCM specifies the number of times that the pixel with value i occurred horizontally adjacent to a pixel with value j. The resulting matrix was analyzed and based on the existing information, the feature vectors are formed [14].

$$CONTRAST = \sum_{n=0}^{G-1} n^2 \left\{ \sum_{i=1}^{G} \sum_{j=1}^{G} P(i,j) \right\}, \quad |i-j| = n \qquad (1)$$

HOMOGENEITY:

$$= \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1+(i-j)^2} P(i,j) \qquad (2)$$

$$ENTROPY = -\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i,j) \times log(P(i,j)) \qquad (3)$$

$$VARIANCE = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i-\mu)^2 P(i,j) \qquad (4)$$

$$SHADE = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i+j-\mu_x-\mu_y\}^3 \times P(i,j) \qquad (5)$$

$$PROM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i+j-\mu_x-\mu_y\}^4 \times P(i,j) \qquad (6)$$

Based on colour, the feature vector is computed directly from RGB colour space [15] and it is given by:
FV=[FE(R), FE(RG), FE(RB), FE(G), FE(GB), FE(GR), FE(B), FE(BG), FE(BR)]
Where FE represents creating a GLCM and computing homogeneity for this matrix .

3.1 CLASSIFICATION USING SUPPORT VECTOR MACHINE (SVM):

The standard SVM is a binary classifier which has found widespread use in pattern recognition problems such as image and audio recognition, handwriting recognition, medicine, science, finance and so on.

The support vector machine or SVM framework is currently the most popular approach for "off-the-shelf" supervised learning.

There are three properties that make SVMs attractive:

1. SVMs construct a maximum margin separator—a decision boundary with the largest possible distance to example points. This helps them generalize well.
2. SVMs create a linear separating hyper plane, but they have the ability to embed the data into a higher-dimensional space, using the so-called kernel trick. Often, data that are not linearly separable in the original input space are easily separable in the higher dimensional space. The high-dimensional linear separator is actually nonlinear in the original space. This means the hypothesis space is greatly expanded over methods that *use* strictly linear representations.
3. SVMs are a nonparametric method—they retain training examples and potentially need to store them all. On the other hand, in practice they often end up retaining only a small fraction of the number of examples—sometimes as few as a small constant time the number of dimensions. Thus SVMs combine the advantages of nonparametric and parametric models: they have the flexibility to represent complex functions, but they are resistant to over fitting. The input points are mapped to a high dimensional feature space, where a separating hyper-plane can be found. The algorithm is chosen in such a way as to maximize the distance from the closest patterns, a quantity which is called the margin. SVMs are learning systems designed to automatically trade-off accuracy and complexity by minimizing an upper bound on the generalization error. In a variety of classification problems, SVMs have shown a performance which can reduce training and testing errors, thereby obtaining a higher recognition accuracy. SVMs can be applied to very high dimensional data without changing their formulation.

IV.   Results and Discussion

Figure 4.1 Abnormal Image

Fig 4.1 shows the classification result using Support Vector Machine (SVM) classifier for abnormal image. The preprocessed retinal image is converted to L*a*b Color space. The color component alone is extracted. It is then applied to the K-means Clustering algorithm. The result consists of five clusters. Since Optic Disc and Exudates are homogenous in color, cluster containing Optic Disc is selected for feature extraction. Based on the feature extracted, the SVM is trained for normal and abnormal images. Finally, image is classified as exudates or non-exudates using SVM.
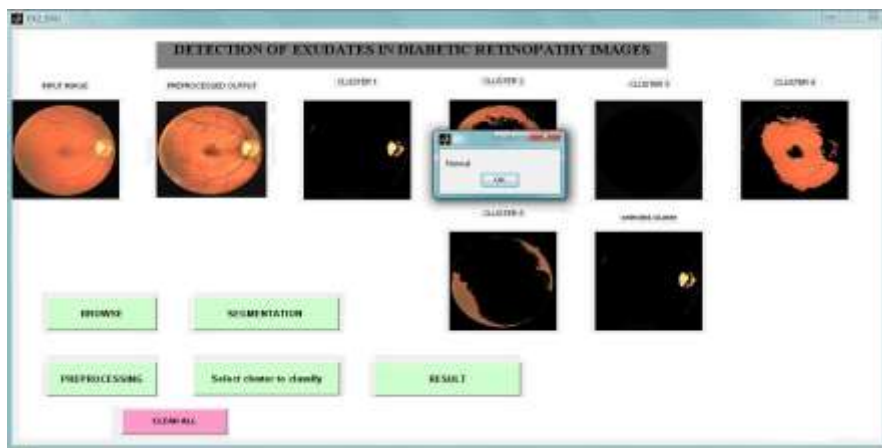

Figure 4.2 Normal Image

Figure 4.2 shows the classification result using Support Vector Machine (SVM) classifier for normal image.

I. Feature Extraction from DRIVE images

Image	Mean	SD	Energy	Dissimilarity
Sample 1	83.361	55.696	0.0305	0.0388
Sample 2	83.041	53.928	0.0297	0.04
Sample 3	86.734	55.374	0.0321	0.0418
Sample 4	57.007	35.337	0.0136	0.0277
Sample 5	93.885	62.352	0.0385	0.045

II. Feature Extraction from Real Time Images

Image	Mean	SD	Energy	Dissimilarity
Sample 1	25.55	26.98	0.0002	0.0231
Sample 2	75.10	56.52	0.0010	0.0852
Sample 3	43.04	36.96	0.0004	0.0326
Sample 4	39.07	31.19	0.0003	0.0459
Sample 5	36.73	31.49	0.0003	0.0372

Table I shows the feature extraction of five samples like mean, standard deviation, Energy and Dissimilarity has been extracted from a publicly available DRIVE database. Table II shows that the feature extraction has been done for five samples of real time images.

III. Feature Extraction of Normal Image Samples (N1-N5) and Abnormal Image samples (E1 – E10)

Images	Correlation	Contrast	Energy	Homogeneity
E1	0.1437	0.0076	0.9801	0.9961
E2	0.1766	0.0145	0.9701	0.9927
E3	0.1381	0.0231	0.9533	0.9884
E4	0.3059	0.0145	0.9705	0.9927
E5	0.3751	0.0209	0.9573	0.9895
E6	0.1637	0.0209	0.9768	0.9946
E7	0.2153	0.0092	0.9807	0.9953
E8	0.2368	0.0077	0.9807	0.9961
E9	0.1780	0.0255	0.9838	0.9872
E10	0.1593	0.0223	0.9463	0.9888
N1	0.1037	0.0252	0.9527	0.9873
N2	0.1381	0.0323	0.9484	0.9838
N3	0.0981	0.0199	0.9324	0.9900
N4	0.1201	0.0231	0.9614	0.9909
N5	0.1161	0.0089	0.9112	0.9211

IV. Feature Extraction of Normal and Abnormal Image with single reference Image

PARAMETERS	NORMAL IMAGE VALUES	ABNORMAL IMAGE VALUES	PARAMETERS	NORMAL IMAGE VALUES	ABNORMAL IMAGE VALUES
CONTRAST	6.9584	1.5954	VARIANCE	1.7495	2.5904
CORRELATION	9.6409	7.1263	SUM ENTROPY	7.0787	9.528
ENTROPY	8.265	9.9922	DIFFERENCE ENTROPY	1.3115	3.1583
HOMOGENEITY	9.7966	9.508	DISSIMILARITY	1.1345	2.682
CLUSTER SHADE	6.8527	-1.1504	INVERSE DIFFERENCE	8.709	9.7816
MAXIMUM PROBABILITY	4.6249	6.052	CLUSTER PROMINANCE	1.1015	1.6666

Table III shows the feature extraction of 10 abnormal images from E1 to E10 and 5 normal images from N1 to N5. The correlation, contrast, energy and homogeneity has been compared for abnormal and normal image. Table IV shows different parameter values of a single normal and a single abnormal image.

V. CONCLUSION

Exudates are one of the earlier signs of diabetic retinopathy. The low contrast digital image is enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE). The Contrast enhanced color image is segmented using K-means clustering, which is one of the simplest unsupervised learning algorithm for image segmentation. The diabetic retinopathy images were collected from nearby hospital and the features extracted have been compared with publicly available STARE and DRIVE database. K-means clustering takes less computational time compared to FCM. It provides more color information from which the result of classification will be improved. To Classify these segmented image into exudates and non-Exudates, a set of features based on texture and color are extracted using Gray Level Co-Occurance Matrix (GLCM). The selected features are classified into exudates and non-exudates using Support Vector Machine (SVM) Classifiers.. Using this approach, the exudates are detected with 96% success rate. The Detection of Microaneurysm which is one of the earlier symptoms of Diabetic Retinopathy can be predicted and its performance can be compared in the future work. Using the same method detection of maculopathy can be done in future and the features can be extracted.

REFERENCES

Journal Papers:
[1] Gwenole Quellec, Stephen R. Russell, and Michael D. Abramoff, Senior Member, IEEE"Optimal Filter Framework for Automated,Instantaneous Detection of Lesions in Retinal Images" *IEEE Transactions on medical imaging,* vol. 30, no. 2,pp. 523-533,February 2011.
[2] Alireza Osareh, Bita Shadgar, and 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,pp.535-545,July 2009.
[3] Carla Agurto, Eduardo Barriga, Sergio Murillo, Marios Pattichis, Herbert Davis, Stephen Russell, Michael Abrrmoff, and Peter Soliz "Multiscale AM-FM Methods for Diabetic Retinopathy Lesion Detection". *IEEE Trans. Medical. Imaging* Vol. 29, No. 2,pp.502-512, February 2010.
[4] T. Walter, J. Klein, P. Massin, and A. Erginary,. "A contribution of image processing to the diagnosis of diabetic retinopathy, detection of exudates in colour fundus images of the human retina". *IEEE Trans. Medical. Imaging*,Vol. 21, No. 10, pp.1236–1243, October. 2002.
[5] Akara Sopharak , Bunyarit Uyyanonvara and Sarah Barman **"**Automatic Exudate Detection from Non-dilated Diabetic Retinopathy Retinal Images Using Fuzzy C-means Clustering".Journal of *sensors/2009.* ISSN 1424-8220.www.mdpi.com/journal/sensors.
[6] Niemeijer, B.V Ginnekan S.R, Russell. M and M.D. Abramoff "Automated detection and differentiation of drusen, exudates and Cotton wool spots in digital color fundus photographs for diabetic retinopathy diagnosis ", Invest. Ophthalmol Vis. Sci., Vol.48,pp. 2260-2267, 2007.
[7] Niemeijer.M, Abramoff.M.D, Van Ginneken.B, "Information fusion for Diabetic Retinopathy CAD in Digital color fundus photographs" *IEEE Transactions on medical imaging,* vo7l. 26, no. 10,pp. 1357-1365, October, 2007.
[8] Ricci.E, Perfetti.R, "Retinal Blood vessel segmentation using Line operators and Support Vector Classification" *IEEE Transactions on medical imaging,* vol. 28, no. 5,pp. 775-785, March 2009.
[9] Goatman.K.A, Fleming. A.D, Philip. S, William. G.T, "Detection of New vessels on the Optic Disc using Retinal photographs" *IEEE Transactions on medical imaging,* vol. 30, no. 4,pp. 972-979, April 2011.
[10] Deepak.K.S, Sivaswamy. J, "Automatic assessment of macular edema from color retinal images" *IEEE Transactions on medical imaging,* vol. 31, no. 3,pp. 766-776, March 2012.
[11] Huiqili, Chutatape. O, "Automated feature extraction in color retinal images by a model based approach" *IEEE Transactions on Bio-Medical Engineering,* vol. 51, no. 2,pp. 246-254, February 2004..
[12] Aquine. A, Gegundez, Aries. M.E, Marin.D, "Detecting the Optic Disc boundary in digital fundus images using morphological, edge detection and feature extraction technique" *IEEE Transactions on medical imaging,* vol. 29, no. 11,pp. 1860-1869, November 2011.
[13] Tobin.K.N, Chaum.E, Govindasamy.V.P, "Detection of anatomic structures in human retinal imagery" *IEEE Transactions on medical imaging,* vol. 26, no. 12,pp. 1729-1739, December 2007.
[14] Akara Sopharak, Mathew N. Dailey, Bunyarit Uyyanonvara, Sarah Barman, Tom Williamson,Yin Aye Moe, "Machine Learning approach to automatic Exudates detection in retinal images from diabetic patients", Journal of Modern optics,2009.
[15] Fleming. AD, Philips. S, Goatman. KA, Williams. GJ, Olson. JA, sharp. PF, "Automated detection of exudates for Diabetic Retinopathy Screening", Journal on Phys Med.and Bio., vol. 52, no. 24, pp. 7385-7396, 2007.
Proceedings Papers:
[16] Doaa Youssef, Nahed Solouma, Amr El-dib, Mai Mabrouk, "New Feature-Based Detection of Blood Vessels and Exudates in Color Fundus Images"IEEE *conference on Image Processing Theory, Tools and Applications*,2010,vol.16,pp.294-299.
[17] Sanchez. C.I, Mayo.A, Garcia. M, Lopez.M.I, Hornero. R, "Automatic Image processing Algorithm to detect hard exudates based on Mixture models" IEEE *conference on Engineering in medicine and Biology society*, pp. 4453-4456, September 2006.
[18] Pradeep Kumar. A. V, Prashanth. C, Kavitha.G, "Segmentation and grading of Diabetic retinopathic exudates using error boost feature selection method "*World Congress on Information and Communication Technologies,* pp. 518-523, December 2011.
[19] C. Sinthanayothin, "Image analysis for automatic diagnosis of Diabetic Retinopathy", *World Congress on Information and Communication Technologies,* pp. 522-532, December 2000.