Early Detection and Classification of Diabetic Retinopathy using K-Means Colour Compression and Fuzzy Logic

Sumandeep Kaur* and Daljit Singh**
*M.Tech. Student, Department of ECE, GNDEC, Ludhiana-141006 India
sumansaraon44@gmail.com
**Assistant Professor, Department of ECE, GNDEC, Ludhiana-141006 India
daljithathan@gmail.com

Abstract: Diabetic retinopathy is anailmentinitiateddue to diabetes mellitus. The longer an individual is suffering from diabetes the more is the possibility of diabetic retinopathy. There are certain abnormalities that are caused by DR, like Micro-aneurisms, Cotton wool spots, soft exudates and hard exudates. The cotton & soft exudates are the most primitiveindication of DR while Hard exudates &Hemorrhages are experienced in the advancedphases of the diabetic retinopathy. The aim of the projected work is to effectively diagnose and categorize diabetic retinopathy. K-means compression technique is helpful in decreasingcolor dimensions by clustering the fundus image in various areas of concern. These areas are then segmented out and analyzedover various region characteristics. Lastly, theclassification of fundus imagesis done with the help of knowledge-based (KB) fuzzy inference system (FIS) by using the effective characteristics of the region. The accuracyand sensitivity of the system arecalculatedto be 94.7% and 96% respectively.

Keywords: Diabetic retinopathy fuzzy inference system, and k-means clustering.

Introduction

Diabetes is a chronic syndrome that arisesdue to a decrease in production of insulin, or alternatively, due to ineffective use of produced insulin by the body. Insulin is a hormone that is mainly used forcontrolling blood glucose level. Diabetes led to the weakness of the blood vessels in the body. The small and delicate blood vessels are more prone to this weakness. The weakness in retinal blood vessels,lead to structural variationswithin the retina, which is known as diabetic retinopathy. The blood vessels in retina undergo many changes like swelling, leakage or may close off completely. Other retinal changes may include the development of new abnormal blood vessels on the retina. Such changes evolve from one phase to another [2].

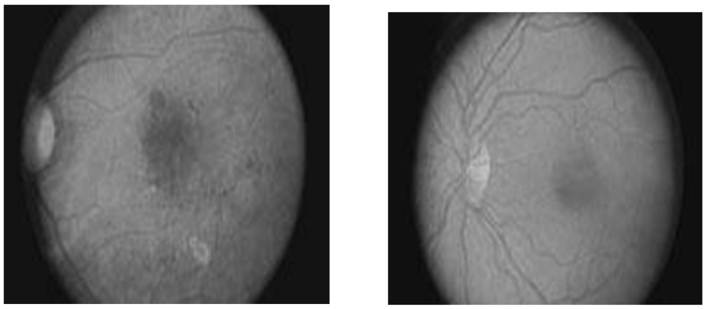

Figure 1: Affected Diabetic eye (left) and Normal eye(right)

WHO reported that in the year 2012, approximately 347 million people around the world havediabetes. WHO projected that diabetes may become the 7^{th} leading root cause of death by 2030[5]. WHO also predicted that diabetic population in India will rise to 79.4 million by the year 2030. This digit is largest among any nation in the world. There are very less number of ophthalmologists in India for the treatment of DR.

Diabetic retinopathy usually impacts both the eyes. It remains undetected at early stages, but at later stages, it may lead to total vision loss that can't be reversed.Keeping strict control of blood glucose leveland regular eye screening by a doctor are means to avoid diabetic retinopathy and vision loss. The aim of this algorithm is to reduce the manual workindiagnosingthe diabetic retinopathy. Specialized training of ophthalmologists is required to prevent diabetic retinopathy and blindness in diabetic patients.

Literature Studies

Md. Jahiruzzaman and A. Hossain [8] offered a diabetic retinopathy diagnosis approach centred on several exudates and detection of hemorrhage. The sensitivity and precision of the detector were achieved to be 98.2% and 92.3% respectively using a minor computation spell as compared to retinal vessel centred diagnosis of diabetic retinopathy. The offered classification using fuzzy system had advanced performance having accurateness upto 96.67%. M. Rajput, et al [7] exposed lesions of non-proliferative diabetic retinopathy via wavelets and categorized it by K-means clustering. The suggested algorithm was verified databases that are available online such as DRIVE, DiarectDBO, STARE, and SASWADE. K-means color compression was employed on these datasets and accurateness of 95% was attained in categorizing NPDR. M Kumar, et al [11] employed the connection of an abnormal wideness within the vessels of blood to discover an estimated position of an optic disk. Some images of the database with diverse contrast, illumination, and DR stages were assessed by algorithm had provided success rate of 90% for hemorrhages. Additionally, the distinct features extracted had provided about 95% of success rate for a microaneurysm, sensitivity of 95% and specificity of 94% for identification of exudates and an accuracy of 97% for optic disk localization. R. Manjula and V. Rajesh [4] established an algorithm for Microaneurysmsrecognition centered on the study of Eigenvalue by the use ofthe hessian matrix in the retinal image. Eigenvalue analysis was operative in the recognitionof dark lesion in the retinal Fundus image. This method was estimated by using 89 images from a database, the true positive level of visible MAs is 91 % with eight wrong positive images.S Kasurde and S.N. Randive [15] had completed an involuntary diagnosis of Proliferative Diabetic Retinopathy. These morphologicalprocesses were engaged for manifold orientations like 45°, 90°, 135° and 180°. Feature extraction was finished by windowing images into 50x50 to estimate the count of pixels in a vessel of each window. In case if these numbers were superior to the threshold value then PDR was identified.

Present work

Steps followed in Present Work

In this work, a method for early diagnosis as well as classification of diabetic retinopathy by means of fuzzy logic and k-means clustering is used. The diagnosis of diabetic retinopathy is completed using the finding of various exudates and hemorrhage. Initially, Sobel edge detector is employed for detecting the edges of the eyeball. Then k-means color compression is employed to decreasecolor dimensions for segmenting exudates by using histogramcenteredthresholding with a lesser computational load.

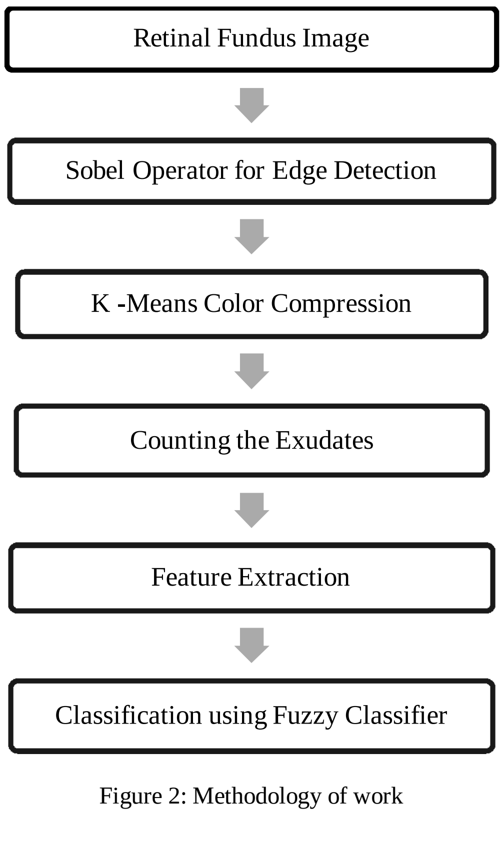

Figure 2: Methodology of work

Sobel Edge Detection

Edge detection of an image is a way of finding edges of the image which is significant in discovering the approximate absolute gradient scale at every point of agrayscale image at an input. The Sobel operation undertakes 2-D spatial gradient calculations on the images. Converting a 2-D array of pixels to the statistically uncorrelated record sets improves the deductionin the amount of redundantdata that results in, declination in the total data required for signifying a digital image [8].

A Sobel operator is used for isolating eyeball from its background by firstly perceiving edges of the eyeball. It is gradient centered edge detection technique. Gradient (G_x, G_y) can be measured by measuring partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ at each pixel location. Hence, the gradient of an image is specified as:

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

Mostly detection of edges is based on the supposition that the edgesexistwherever discontinuity in an intensity function exists or wherever a sharp intensity gradient is present in an image. Then derivative of an intensity value is taken across an image and pointswere discovered at which derivative is highest. Using this value the edges can be founded.

Gradient is a vector quantity. The rate of change of pixel value with respect to distance from x and y-axis is dependent upon the constituents of this gradient vector. Edge detector using Sobelutilizesthe pair of 3x3 convolution covers, one approximating gradient along x-direction whileanother approximating gradient along y-direction. This indicator is tremendously profound to noise in an image and efficiently highlightsit as edges.

Sobel operationcan be measuredby shortening the partial derivative if the 3x3 kernel is known. Sobel operator is given as:

$$G = |G_x| + |G_y| = \sqrt{{G_x}^2 + {G_y}^2}$$

$$= |l(3,1) + 2l(3,2) + l(3,3)| - |l(1,1) + 2l(1,2) + l(1,3)| + |l(1,3) + 2l(2,3) + l(3,3)| - |l(1,1) + 2l(2,1) + l(3,1)|$$

here, l stands for the pixel intensity ofthe kernel matrix. Hence edges could be found by moving and convolving a kernel throughout the image.

K-means clusteringfor detection of exudates

The clustering method is employed to place alike data objects intoonecluster. K-means is a widespread grouping practice. This is an unsupervised process for organizing the datasets at input into numerous classes grounded on the characteristic distance. Here, for picking the color linked to the exudates, k-means is employed for colorcentered segmentation of the retinal image. Image colorsare stated as three-dimensional vector i.e. RGB of the pixel intensity. Like colors could be clustered together while some amount of constituents could be reducedto neglect redundancy in components during compression [7].

K-Means clustering anticipates to paneling n substances into k clusters in which every article is appropriate to a cluster with an adjacent mean. This scheme harvests exactly k altered clusters of maximumachievabledissimilarity.There are k clusters that are not given previously, lead to extreme separation and this value is to be figured from the given data [8]. To diminish the total intra-cluster variance is an actual objective of k-means color compression. It is arepetitive procedure and twitches with a preliminary centroids estimation. The clusters are designed contingent on adjacent centroids. This procedure redefines centroids repetitively till an error congregates. If $m_1,\ldots,m_k$ is an initial set of k means, then the adjoining means can be measured by a manipulative squared Euclidean distance which can be represented as:

$$S_i^{(t)} = \left\{x_p : \left\|x_p - m_i^{(t)}\right\|^2 \le \left\|x_p - m_j^{(t)}\right\|^2 \ \forall j, 1 \le j \le k\right\}$$

The centroid of new clusters is used to measure the new mean. Following is the formula for measuring the new mean:

$$m_i^{(t+1)} = \frac{1}{\left|S_i^{(t)}\right|} \sum_{x_j \epsilon S_i^{(t)}} x_j$$

As exudates are much similar to the red constituent of an image, so the color compressed image is firstly reduced to the red subpart. For this image, the histogram is then drawn and this histogram is used for selecting superior and inferior thresholds. Using these threshold values, exudates can be segmented by computing the binary image. At next step,a morphological operation is employed to the segmented exudates. Lastly, a sectioncentered bounding box is castoff to recognize as well as to compute the exudates.

Classification by using Fuzzy Inference System

Fuzzy inference is the procedure of using a fuzzy logic for the mapping from input to output. This procedure comprises ofthe membership function, if-then rules, and fuzzy operations. FISiseffectively applied in various fields like automatic controller, classification of data, decision-making, skilledorganizations, and PCvisualization. The DR calculation in thisscheme is done by selectingthe actual characteristics. The segmented exudates include section assets characteristics. Here, a fuzzy inference system is employed and developed with this operative attributes. Thisfuzzy inference system is intended by3 classes i.e. normal, medium, and severe.

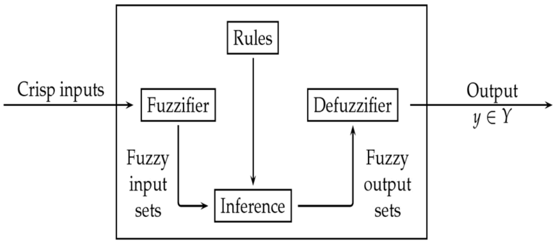

Figure 3: Fuzzy inference systems

The working procedure of FIS involves the steps given below:

1. Equate input variable with membership functionofan antecedent portion to attain membership valueatevery linguistic mark.
2. To obtain the firing strength of every rule, membership values on premise portion are combined
3. Generate the qualified components (either crisp or fuzzy) or every rule based on the firing strength.
4. To attain crisp output, the qualified components are aggregated.

Results and Sensitivity Analysis

DIARETDB0 database is used in this experimentation for identifying normal plus abnormal retinal images.This database comprises of colored fundus images of both normal as well asthe abnormal diabetic eye. The images were captured at Kuopio Universityand hospital with digital fundus camera with 50-degree field-of-view.

Detection of Diabetic Retinopathy

Retinal fundus image with diabetic retinopathy is applied as an input to the detector system.

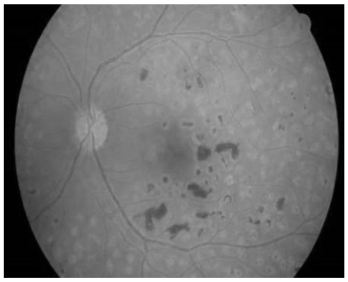

Figure 4:Abnormalimage affected by diabetic retinopathy

Then Sobel operation is used for an edge detection of the image. For this, first of all,an RGB image is converted into gray and then Sobel operation is applied.

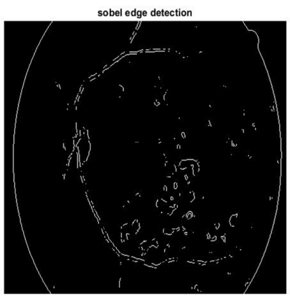

Figure 5: Sobel edge detection

K-means color compression is then applied onthe image to color compress the image.The color of an imagethat is separated from its background is compressed in 30 colors ofan RGB image. The figure shows the compressed image after segmentation.

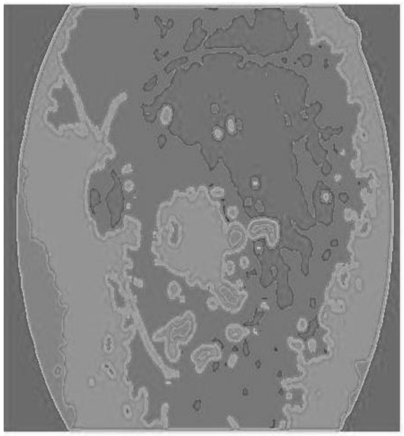

Figure 6: Color image segmentation

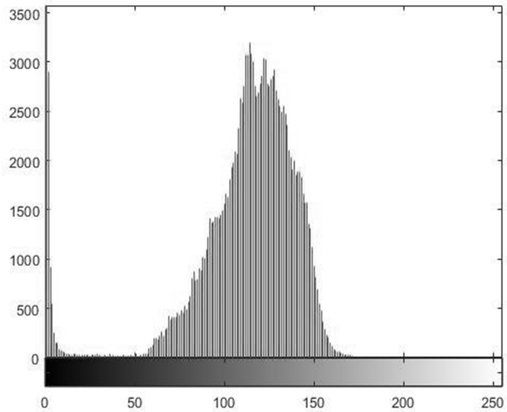

Figure 7: Histogram of thecolor compressed image

The threshold value is selected from the histogram of color compressed image and further this threshold is used to construct a binary image. Exudates are segmented from the image by filling holes and by performing morphological operations on abinary image. The structuring components are employed for improved visualization.

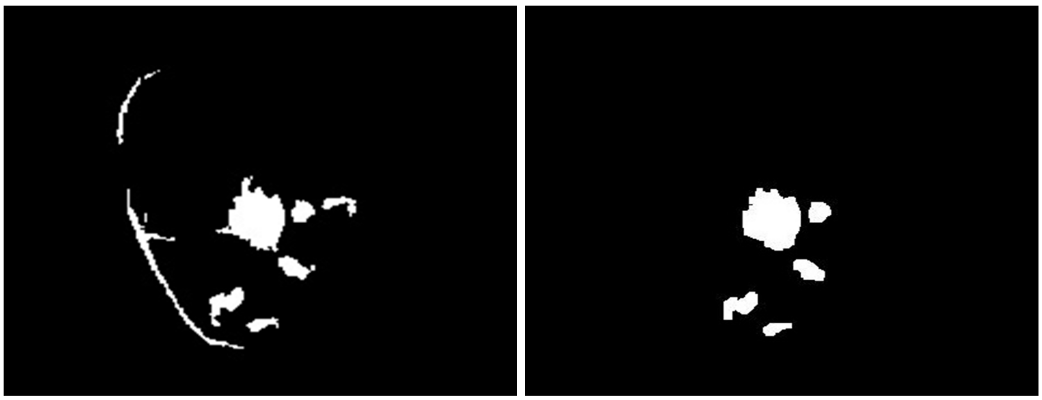

Figure 8: Segmented exudates earlier (left) and after the filling (right) of holes

For identification and enumeration of segmented exudates, region properties function i.e. bounding box is used. For better visualization, an original image is merged with the segmented imagea having bounding box for representing detected exudates.

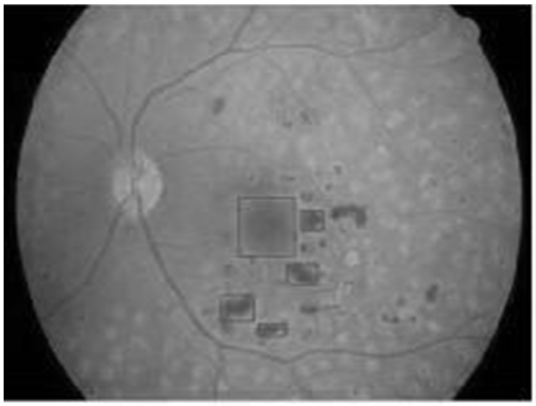

Figure 9: Merged image showing detected exudates

Sensitivity and accuracy can be calculated to measure the performance of the detector system. The confusion matrix for the system is shown in the table below:

Table 1: Confusion Matrix

		Predicted data		Total
		True	False	
Real data	True	TP	FN	T
	False	FP	TN	P
Total		T'	P'	T+P or T'+P'

In this table, TP, FP, TN, FN represents the total of true positive, false positive, true negative and false negative trials respectively. TN and TPindicatethatsystem is receiving right results, whereas FN and FPindicatethat system is receiving wrong results. P represents no. of elements ina positive set while N represents no. of elements in a negative set.

Sensitivity: To measure true positive ratio of classified images

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

Accuracy:To measure percentage of correctly classified images

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP}$$

Classification of Diabetic Retinopathy

Various regions of interest are obtained by segmentation ofimage after color compression. Dot exudates lie in the lower intensity range of a red portion of the RGB image. The hard exudates lie on higher intensity range while soft exudates lie on threshold region. For segmentation of these regions, region properties functions such as area, bounding box etc are used. The system is classified into three levels i.e. normal, moderate and severe.

Table 2: Fuzzy set and its numeric values

Input variable	Normal	Moderate	Severe
Cotton and Soft Exudates	0-15	10-60	50-210
Hemorrhage and Hard Exudates	0-5	3-20	16-40
Area	0-500	350-2100	1600-7000

Table 3: Fuzzy Rules for Classifier System

Rule	Cotton& Soft Exudates	Hemorrhage& Hard Exudates	Areas	Class
Rule 1	N	N	N	1
Rule 2	M	N	N	1
Rule 3	M	N	M	2
Rule 4	M	N	S	2
Rule 5	S	N	M	2
Rule 6	S	N	N	2
Rule 7	S	N	S	3
Rule 8	N	M	N	2
Rule 9	N	M	M	2
Rule 10	N	M	S	3
Rule 11	M	M	N	2
Rule 12	M	M	M	3
Rule 13	M	M	S	3
Rule 14	S	M	N	2
Rule 15	S	M	M	3
Rule 16	S	M	S	3
Rule 17	N	S	N	3
Rule 18	N	S	M	3
Rule 19	N	S	S	3
Rule 20	M	S	N	3
Rule 21	M	S	M	3
Rule 22	M	S	S	3
Rule 23	S	S	N	3
Rule 24	S	S	M	3
Rule 25	S	S	S	3

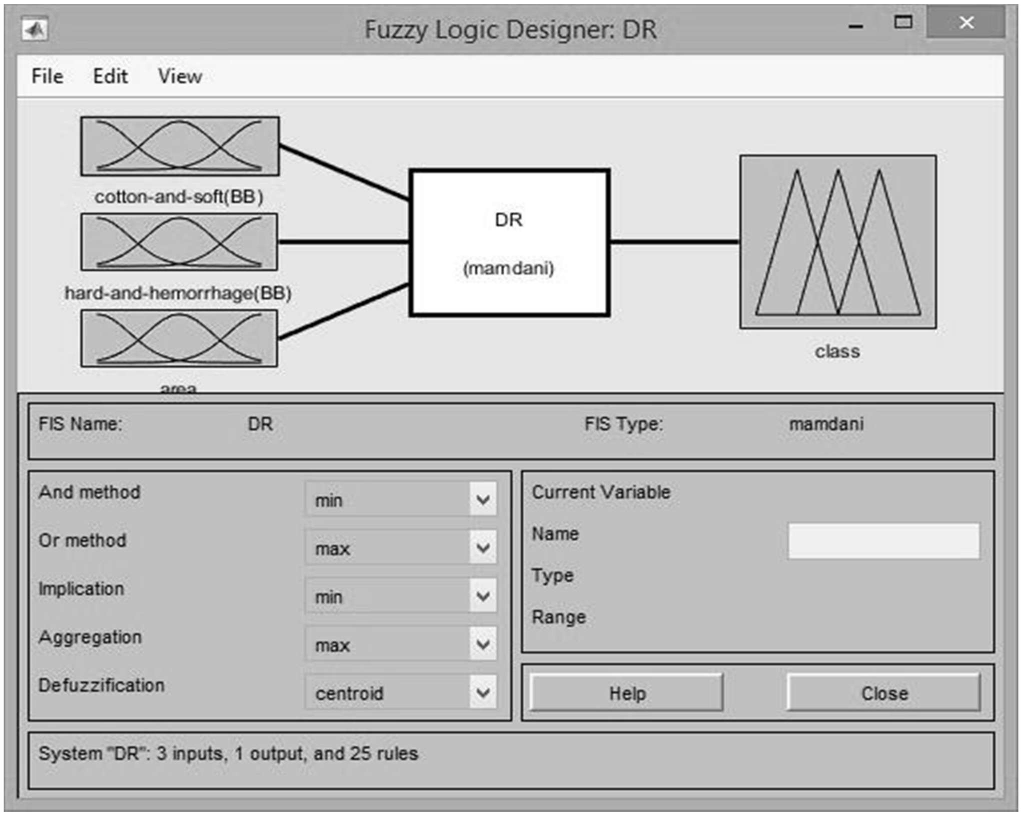

Figure 10: Mamdani fuzzy logic

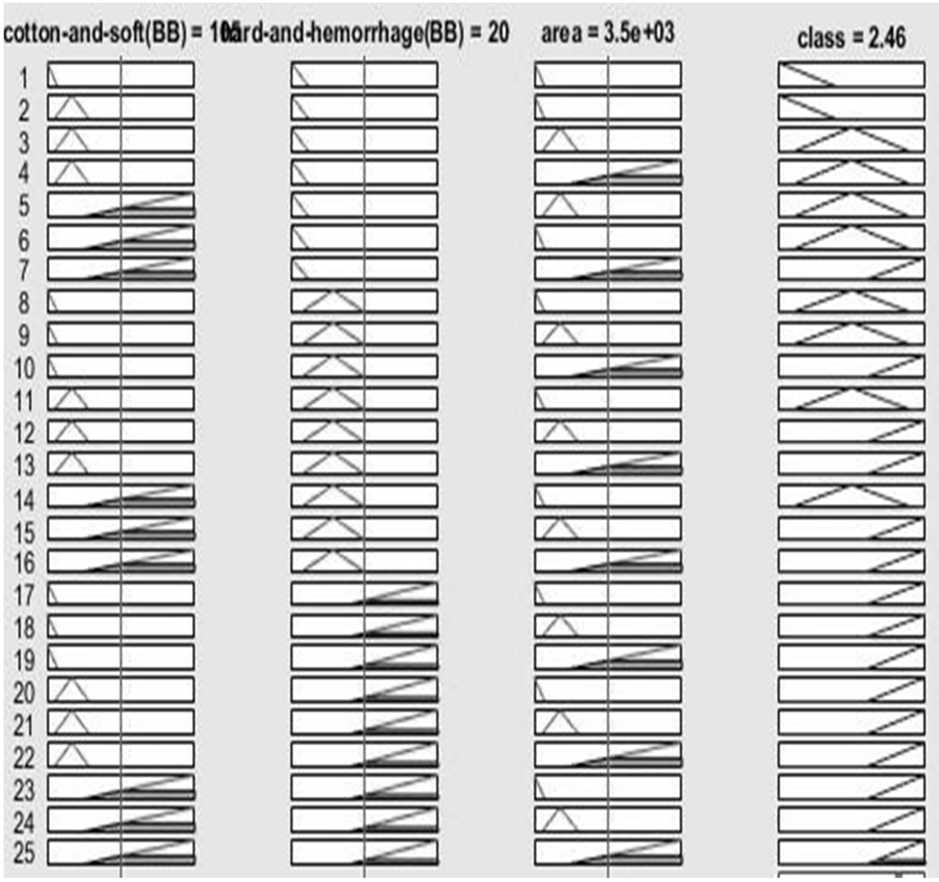

Figure 11: Rule Viewer

Class 1 signifies the primary stage, class 2 signifies intermediate stage and class 3 signifies severe stage of DR. The FIS system is applied to 25 images which give the outcome after classification.

Calculations

In this study total of 25 images were takenout for which 7 images were of a normal eye having no signs of diabetic retinopathy while 18 images were of a diabetic eye having cotton soft and hard hemorrhages. After performing the operation total of 24 images were identified accurately while 1 image shows normal class while it was in the intermediate class. Therefore 24 out of 25 images were detected accurately.

Therefore, TP = 18

TN = 6

FP = 1

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} = \frac{18 + 6}{18 + 0 + 6 + 1} = \frac{24}{25} = 96\%$$

$$Sensitivity = \frac{TP}{TP + FN} = \frac{18}{18 + 1} = 94.7\%$$

Therefore, accuracy and sensitivity of the system is attained as 96% and 94.7% respectively.

Conclusions

In this work, a methodology for initial diagnosis and for classifying diabetic retinopathy via fuzzy inference system and k-means clustering is developed. Diabetic retinopathy diagnosis is done using detection of different exudates and hemorrhage. Firstly Sobel edge detector is used for identifying the edges of the eyeball. Then k-means clustering method is appliedfor lessening the color dimensionsto segment exudates by employing the histogram based thresholdingusing adecreased computation load. The accuracy and sensitivity of system are attained as 96% and 94.7% respectively. This system could be employed in hospitals for the early diagnosis of the diabetic retinopathy to prevent vision loss.

Future scope

The operative screening of fundus images of Diabetic affected population for recognizing and classifying various phases of Diabetic Retinopathy could decrease the probability of unexpected loss of vision in diabetic population. As the prevailing schemes are relatively slower in a procedure, animmediate application of detection process coulddeliver the superior performance. Aiming at the execution on immediate platforms, DSP practices should be chosen as these couldoffer aninnovativeoutcomeinrecognizing the Diabetic Retinopathy[5]. Such screening systems for the recognition of diversephases of Diabetic Retinopathy are primarily beneficial to affected peoplein rural zonesthatstay ignorant of the existence of Diabetic Retinopathy. Further accuracy and sensitivity of the system can be enhanced by selecting better segmentation and classification techniques.

References

[1] T. Walter et al., “A contribution of image processing to the diagnosis of diabetic retinopathy, detection of exudates in color fundus images of the human retina,” in IEEE Trans. Med. Imag., vol. 21, pp. 1236–1243, Oct. 2002.

[2] M. B. Patwari et al., "Personal Identification algorithm based on Retinal Blood Vessels Bifurcation", in Int. Conf. on Intell. Comput.Applicat., 2014©IEEE, doi: 10.1109/ICICA. 2014.51.

[3] Dr.Vijayachitra. et al. "Analysis Of Diabetic Retinopathy Images Using Blood Vessel Extraction," Int. J. of Adv. Eng. Research and Studies, vol. 1, issue 2, pp. 89-91, Mar. 2012.

[4] R. Manjula and V. Rajesh, “Early detection of diabetic retinopathy from retinal fundus images using Eigen value analysis,” in Int. Conf. on Control, Instrum.,Commun. andComput. Tech., vol. 1, no. 15, pp. 347-351, 2015.

[5] World Health Organization, “Global Health Estimates: Deaths by Cause, Age, Sex and Country 2000-2012”, in Department of Health Statistics and Information Systems WHO, Geneva , 2014.

[6] A.D. Fleming et al., “Automated detection of exudates for diabetic retinopathy screening,” in Phys. Med. Biol., vol. 52, pp. 7385–7396, 2007.

[7] M. Rajput et al., “Detection of Non-Proliferative Diabetic Retinopathy Lesions using Wavelet and Classification using K-Means Clustering,” in 2015 Int. Conf. on Commun. Networks, pp. 381- 387, 2015.

[8] Md. Jahiruzzaman and A. Hossain, “Detection and Classification of Diabetic Retinopathy Using K-Means Clustering and Fuzzy Logic,” in 18th Int. Conf. on Comput. And Inform. Techno.,pp 534-538, Dec 2015.

[9] Usman M. et al.,"Automated Detection of Dark and Bright Lesions in Retinal Images for Early Detection of Diabetic Retinopathy," 2011 © Springer, doi: 10.1007/s10916-011-9802-2.

[10] M. B. Patwari et al., "Detection and Counting the Microaneurysms using Image Processing Techniques,” Int. J. of Appl. Inform. Syst., vol. 6, no. 5, pp. 11-17, Oct. 2013.

[11] M. Kumar et al., “Feature extraction from the fundus images for the diagnosis of diabetic retinopathy” in Int. Conf. on Emerging Research in Electron., Comput. Sci. and Tech., vol.1, no. 2, pp. 240-245, 2015.

[12] D. Marin et al., “A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment in variants based features,” in IEEE Trans. on Med. Imaging, vol. 30, no. 1, pp. 146–158, Jan. 2011.

[13] B. Antal and A. Hajdu, "An ensemble-based system for micro aneurysm detection and diabetic retinopathy grading," in IEEE Trans. on Biomed. Eng., vol. 59, no. 6, pp. 1720- 1726, Jun. 2012.

[14] Sopharak et al., “Machine learning approach to automatic exudate detection in retinal images from diabetic patients,” J. of Modern Optics, vol. 57, no. 2, pp. 124–135, 2010.

[15] S. Kasurde and S.N. Randive, “An Automatic Detection of Proliferative Diabetic Retinopathy,” in Int. Conf. on Energy Syst. and Applicat., pp. 135-140, 2015.