ABHIYANTRIKI
An International Journal of Engineering & Technology
(A Peer Reviewed & Indexed Journal)
Vol. 3, No. 5 (May, 2016) http://www.aijet.in/ eISSN: 2394-627X

Automatic Detection of Diabetic Retinopathy using Morphological Operation and Machine Learning

Bhagyashri S. Mankar*
M. Tech. (CSE)
GNIET, Nagpur, Maharashtra, India

Nitin Rout
Assistant Professor
GNIET, Nagpur, Maharashtra, India

Abstract

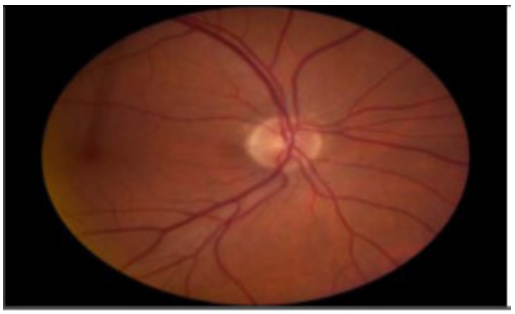
The human eye is an organ which gives a sense of sight. Diabetic Retinopathy is a most common diabetic eye disease which is a leading cause of blindness in India. Diabetic Retinopathy is a disease in which the retinal blood vessels swell and it may even leak. This damages the retina of the eye and may lead to vision loss if the level of diabetes is very high. Early diagnosis of Diabetic Retinopathy can prevent vision loss in patients. The method proposed in this paper for detection of Diabetic Retinopathy disease level emphasizes on determination of two important types of Diabetic Retinopathy; Hemorrhages and Exudates. These types can be extracted using fundus images of patients and processing these fundus images through an appropriate image processing technique. Based on the presence of these types and their amount in the fundus image will determine the level of Diabetic Retinopathy in patients.

Keywords*: Diabetic Retinopathy, Hemorrhages, Exudates, SVM classifier.*

****Author for correspondence*** bhagyashrimankar55@gmail.com

1. Introduction

Diabetic Retinopathy (DR) is an eye disease which occurs due to diabetes. It damages the small blood vessels in the retina resulting in loss of vision. The risk of the disease increases with age and therefore, middle aged and older diabetics are prone to Diabetic Retinopathy. The National Eye Institute estimates that 40 to 45 percent of Americans having diabetes are affected by diabetic retinopathy due to which around 24,000 people become blind every year. Symptoms of diabetes retinopathy do not surface until visual damage to the retina has occurred, usually by partial vision. Therefore regular eye screening is necessary to provide early diagnosis and treatment before significant damage is caused to the retina as it potentially reduce the risk of blindness in these patients by 50%. An early detection of DR enables laser therapy to be performed to prevent or delay visual loss and may be used to encourage improvement in diabetic control. Hence an automatic detection and treatment of the diabetic retinopathy in an early stage can prevent the blindness. Early diagnosis and treatment has been shown to prevent visual loss and blindness. Retinal images obtained by the fundus camera are used to diagnose DR. Automated methods of DR screening help to save time, cost and vision of patients, compared to the manual methods of diagnosis.



Normal Image

Fundus Diabetic Retinopathy Image

Figure 1

2. Review of Literature

Roychowdhury (2014) proposed three stage algorithms for automatic detection and classification. For automated detection, novel two-step hierarchical binary classification is used. For classification purposed GMM, SVM, KNN and ADABOOST methods are used. They take 30 top features like area, variance of Ired channel, Igreen chaneel, I sat of object, major and minor axis length, Mean pixels for Igreen, Ired and Intensity, solidity etc. Gandhi et al. (2013) proposed diagnosis of DR using morphological operations like erosion followed by dilation for detection of exudates. Then GLCM features are calculated. Before that they have used pre-processing operation like colour space conversion, image restoration and enhancement operation. After feature extraction image get segmented and which was applied to SVM and KNN classifier to classify the image according to its severity grade. In this SVM classifier is used to evaluate training data to find a best way to classify images into different cases like mild, moderate or severe retinal images. This diagnosis has been performed on images, which is in .jpg format captured by retinal fundus cameras. And result was matched with that manually outlined by the ophthalmologist. Li (2013) introduced a novel splat feature classification method which was used for the hemorrhage detection in retinal fundus images. The whole retinal colour images are covered and partitioned into non-overlapping segments. Each partition is called as splat which contains pixels with spatial location and same colour. Each splat consists of retinal features which is extracted and are compared with variety of filter bank interactions with neighbouring splats and texture information and shape. Splat features is selected by a filter approach and then apply wrapper approach. Given splats along with their associated feature vectors and reference standard labels, these are passed to classifier for training purpose to detect required objects. Messidor publically available dataset used in this project. Through this 0.96 was achieved for area under the receiver operating characteristic curve at the splat level and also 0.87 was achieved at the image level. Geetha Ramani (2012) proposed a comparative study between two algorithms and compares the result of both algorithms. This comparison was happened between two data mining algorithm i.e. C4.5 decision tree algorithm verses random tree algorithms. In this comparison decision tree gives better result than random tree algorithm. C4.5 algorithm gives 72% accuracy and random tree gives 65% accuracy.

Giancardo et al. (2011) introduced a new methodology for diagnosis of Diabetic Macular Edema (DME) with help of a new set of features which are based on colour wavelet decomposition and automatic lesion segmentation. The single feature vector generated for each image for the OEM diagnosis purpose and thus the feature vector created is based on three types of analysis: Exudates probability map, Color Analysis and Wavelet Analysis. These features are playing a

key role to train a classifier which is able to automatically diagnose DME through the presence of exudation. The accuracy obtained using proposed algorithm is in between 88% to 94%. Agurto et al. (2010) had proposed AM-FM texture feature extraction method which in new than usual method. In usual method segmentation process has performed for feature extraction. In this structure get classified according to type of lesions. And In this accuracy has been calculated according to distance method. This method achieves accuracy up 92% and gives good sensitivity and spasticity. Saleh & Eswaran provided an automated decision-support system for non-proliferative diabetic retinopathy disease based on MAs and HAs detection. In this paper we extract some important features, such as optic disc, fovea, and blood vessels for accurate segmentation of dark spot lesions in the fundus images. Dark object segmentation approach is used to locate abnormal regions such as MAs and HAs. Based on the number and location of MAs and HAs get used to evaluated the extremity level of Diabetic Retinopathy i.e. type of DR. The dataset used consist of 98 colour images to evaluate the performance of the given work. The proposed system achieves 84.31 % and 87.53% values in terms of sensitivity for the detection of MAs and HAs and specificity the system achieves are 93.63% and 95.08% values in terms specificity respectively. In the development of automated screening systems the given system leads to Reliable detection of retinal hemorrhages.

Li Yafen et al. (2013) proposed a new method using different image processing techniques such as image enhancement, morphological image processing and texture analysis. For the classification purposed SVM classifier used. It gives accuracy of 89% then the sensitivity is 90% and specificity is 95%. Proposed paper worked on accuracy of classifier and for that them using Directdb dataset for fundus image. Keerthi Ram mainly focused on clustering based methods to segment the exudates like feature in fundus image. And extract the features were multi space and colour space feature. In this paper mainly focus on processing time factor so due to this its speed is faster than any other techniques. In this it archives accuracy up to 89.6% and gives positive prediction values gives 87%. Osareh (2009) used the c-mean clustering and colour normalization technique for pre-processing operations. In this colour image get segmented using fuzzy and feature of retinal image get extracted using genetic base algorithm. This approach gives 93.6% accuracy and 92.2% sensitivity which are good. A new contextual clustering algorithm has been proposed by JayaKumari & Maruthi (2012) to detect the presence of hard exudates in the fundus images. First the pre-processing stage, then segment the exudates have been done through proposed algorithm. Features extraction is done from the segmented regions which results into the standard deviation, mean, intensity, edge strength and compactness. These extracted features are given as inputs to Echo State Neural Network (ESNN) to differentiate between the normal and pathological image. The dataset consists of a total 50 images have been used to find the exudates. Out of 50, 35 images used to train the ESSN which consist of both normal and abnormal and the remaining 15 images are used to test the neural network. The proposed algorithm achieved 93.0% sensitivity and 100% specificity in terms of exudates based classification.

3. Scope of Study

The project is designed with the fundamental knowledge in digital image processing, basics of statistics, artificial neural network and fundamental mathematics involving matrices. The program is developed using Math works MATLAB software, which it is presented in a Graphical

User Interface. The concepts of Digital Imaging are covered in the following Digital Image, Image Pre-processing, Image Analysis (Texture features), Classification, Recognition.

4. Objectives of Study

The primary aim of this project is to develop a system that will be able to identify patients with PDR and NPDR from colour images obtained from the retina of the patient. These types of images are called fundus images. The different diabetic retinopathy diseases that are of interest include red spots, micro aneurysm and neovascularisation and they fall between PDR and NPDR stages of the disease. In order to perform a task without inserting any command code, simple and user-friendly software is created to facilitate the user when operating the interface. In addition, it benefits the organization by reducing their cost in the long run. As patients in the non-proliferative and proliferative classes are prone to losing their vision, there is a need to identify and notify the affected patients to go for early treatment. Otherwise, the consequences will be irreversible and result in blindness.

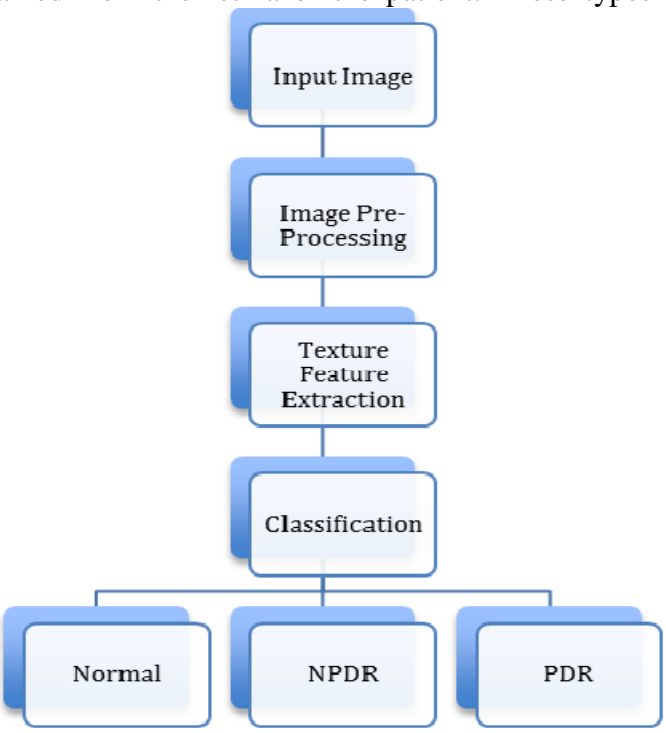


Fig 2: Proposed method development diagram

5. Research Methodology

Firstly this system takes retinal image as a input, after loading the image it will go under pre-processing part where features are extracted from the image and fed these into the classifier for classification of the image as whether this image is normal or having Diabetic Retinopathy.

Image Preprocessing

Image pre-processing is a process to reduce the presence of unwanted features of the image such as noise. The purpose of image pre-processing is to improve the quality of the image being process. As a result, it provides a much accurate results for any image analysis made. Image processing techniques consist of a range of standard image filtering methods such as Median filtering or histogram equalization etc. In this, we are using two image pre-processing methods. Methods are chosen based on the quality of images and the identification of unwanted features that need to be removed to provide a better quality image for analysis.

Median Filtering

Median filtering is an effective method that can remove impulse noise without blurring sharp edges of the image. This is achieved by replacing every pixel value with the median value of the neighbouring pixels. The beneath 3 by 3 window pixel values will illustrate how filtering is done.

20	180	10
25	30	25
15	20	35

Sorting the pixel values in ascending order, we get, 10, 15, 20, 20, 25, 25, 30, 35, 180 the median is the value (25) in the middle position (5th). Therefore, the centre pixel is replaced with the median of 25.

Histogram Equalization

Indicate the given name and family name clearly. Histogram equalization is a method in image processing of contrast adjustment using the image's histogram. This method usually increases the global contrast of many images, especially when the usable data of the image is represented by close contrast values. Through this adjustment, the intensities can be better distributed on the histogram. This allows for areas of lower local contrast to gain a higher contrast. Histogram equalization accomplishes this by effectively spreading out the most frequent intensity values.

6. Feature Extraction

Texture features are derived from the gray level matrix for an image. The following features are selected.

Contrast

Contrast is the measurement of the local variations or differences in the GLCM. It works by measuring how elements do not lie on the main diagonal and returns a measure of the intensity contrast between a pixel and the neighbouring pixels over the whole image. Large contrast reflects large intensity difference in GLCM.

$\sum i \sum j (i-j)^2 P\_d (i,j)$

Homogeneity

Homogeneity measures how close the distribution of elements in the GLCM is to the diagonal of GLCM. Homogeneity weighs values by the inverse of the contrast weight, with weights decreasing exponentially away from the diagonal as shown in equation. The addition of value 1 in the denominator is to prevent the value 0 during division. As homogeneity increases, the contrast typically decreases.

$$\sum_i \sum_j \frac{1}{1+(i-j)^2} P_d(i,j)$$

Entropy

Entropy is understood from the concept of thermodynamics. It is the randomness or the degree of disorder present in the image. The value of entropy is the largest when all elements of the co-occurrence matrix are the same and small when elements are unequal.

$$E_n = \sum_i \sum_j P_d(i,j).1nP_d(i,j)$$

Contrast

Contrast is also known as the second moment of Pd (its moment of inertia about the origin).

$$\sum_{k=0}^{n-1} k^2 P_\delta (k)$$

Entropy

Entropy is smallest when Pd (k) values are unequal and largest when Pd (k) values are equal. Entropy is directly proportional to unpredictability. The above-mentioned features were calculated for d = (0, 1), (1, 1), (1, 0), and the total mean values on the four features were taken.

These features are calculated for various segmented fundus images, i.e. normal fundus images and abnormal (DR) fundus images. These features are applied as input to (Support Vector Machine) SVM classifier.

7. SVM Classifier

A SVM is a discriminative classifier formally defined by a separating hyper plane. In other words, given labelled training data (supervised learning) the algorithm outputs an optimal hyper plane which categorizes examples.

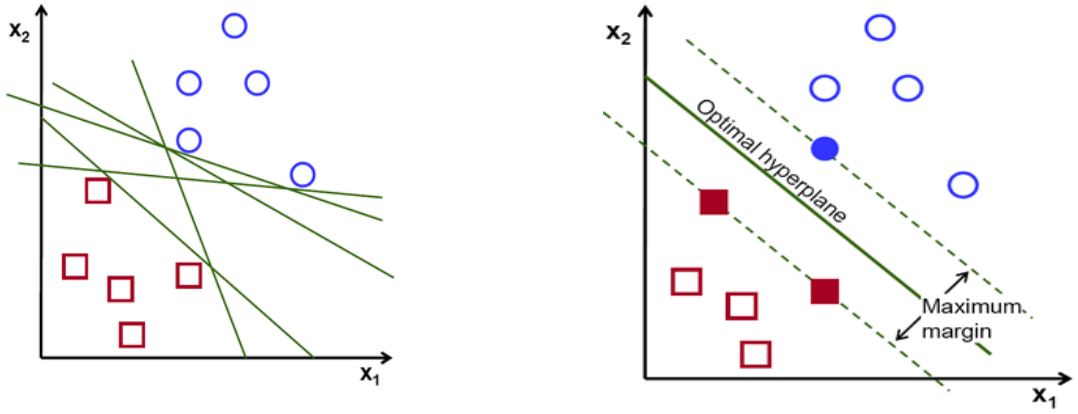


Fig. 3: SVM classifier

Support vector machine training process is applied to analyze training data to find an optimal way to classify images into their respective classes namely PDR, NPDR or Normal. SVM is a robust technique for data classification and regression. SVM models search for a hyper plane that can linearly separate classes of objects shown in Figure. Support vector machine is used to discriminate the various categories. Classification parameters are calculated using support vector machine learning. The training process analyzes training data to find an optimal way to classify images into their respective classes. The training data should be sufficient to be statistically significant. The SVM learning algorithm is applied to produce the classification parameters according to calculated features. The derived classification parameters are used to classify the images. The image content can be discriminated into the various categories in terms of the designed support vector classifier. To fit nonlinear curves to the data, SVM make use of a kernel function to map the data into a different space where a hyper plane can be used to do the separation.

8. Results and Discussion

The Diabetic Retinopathy (DR) level in humans can be detected by scanning the human fundus image for the presence of hemorrhages and Exudates. Macula Edema indicates Mild level of DR, Hemorrhages indicates Moderate Level of DR and Exudates indicate severe level of DR in

humans. The SVM classifier is trained with 100 fundus images which show different levels of DR. After extracting the features, they will go under classification part for identifying that whether the provided input image is normal or having DR. The input test image fed to the classifier appropriately classifies the level of DR based on the training of SVM Classifier.

From Table, it is observed that 75% classification accuracy has been obtained for normal images whereas the DR images gave an accuracy of 68.75% and 89.50% respectively. Thus the abnormality detection is done with higher accuracy.

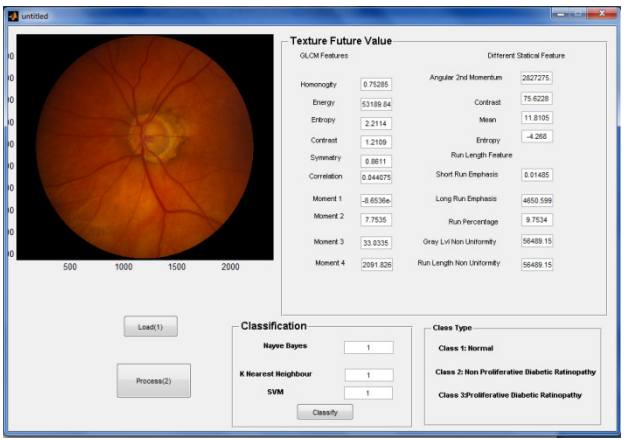


Fig. 4: Final output indicating result of SVM classifier

Table: Result of Support Vector Machine classifier

Class	Number of data for Training	Number of data for testing	Accurately classified	Accuracy
Normal	7	20	15	75%
NPDR	13	16	11	68.75%
PDR	36	19	17	89.50%

9. Conclusion

The method adopted in this paper for early detection of DR disease in humans is reliable and shows accurate results. The method implemented can be used for screening of patients eyeballs for detecting level of DR in a cost effective manner. This technique helps in determining levels of DR in its early stage and thus preventing vision loss. A set of features that describes one case (i.e., a row of predictor values) is called a vector. So the goal of SVM modelling is to find the optimal hyper-plane that separates clusters of vector in such a way that cases with one category of the target variable are on one side of the plane and cases with the other category are on the other size of the plane. The vectors near the hyper-plane are the support vectors. In this paper, SVM classifier is trained with the features of known images, i.e. images whose DR level is already known. This process is known as learning of SVM classifier. The test fundus image is then applied as an input to SVM classifier which provides at the output the level of DR.

References

[1] Roychowdhury, S. (2014). DREAM: Diabetic Retinopathy Analysis Using Machine Learning. *IEEE Journal of Biomedical and Health Informatics*, 18(5).

[2] Gandhi, M. (2013). Diagnosis of Diabetic Retinopathy Using Morphological Process and SVM Classifier. IEEE International conference on Communication and Signal Processing, April 3-5.

[3] Li Yafen. (2013). Automated Identification of Diabetic Retinopathy Stages Using Support Vector Machine. *In:* Proceeding of the 32nd Chinese Control Conference 2013.

[4] GeethaRamani, R. (2012). Data Mining Method of Evaluating Classifier Prediction Accuracy in Retinal Data. *In:* Proceedings of IEEE International Conference on Computational Intelligence and Computing Research.

[5] Rocha, A, Carvalho, T, Jelinek, HF, Goldenstein, S, & Wainer, J. (2012). Points of Interest and Visual Dictionaries for Automatic Retinal Lesion Detection. *IEEE Transactions on Biomedical Engineering*. 59(8), pp.2244-2253.
[6] Atul Kumar, Gaur, AK, Srivastava, M. (2012). A Segment based Technique for detecting Exudate from Retinal Fundus image. *Procedia Technology*. Vol. 6, pp.1-9.
[7] Yazid, H, Arof, H, Isa, HM. (2012). Exudates segmentation using inverse surface adaptive thresholding. *Measurement*. 45(6), pp.1599–1608.
[8] JayaKumari, C, & Maruthi, R. (2012). Detection of Hard Exudates in Color Fundus Images of the Human Retina. *Procedia Engineering*. Vol. 30, pp.297-302.
[9] Giancardo, L et al. (2011). Textureless macula swelling detection with multiple retinal undusimages. *IEEE Trans. Biomed. Eng*. 58(3), pp.795–799.
[10] Agurto, C et al. (2010). Multiscale am-fm methods for diabetic retinopathy lesion detection. *IEEE Transactions on Medical Imaging*. 29(2), pp.502-512.
[11] Osare, A et al. (2009). A Computational Intelligence Based Approach for Detection of Exudates in Diabetic Retinopathy Images. *IEEE Transactions on Information Technology in Biomedicine*. 13(4), pp.535-545.
[12] Sopharak, A et al. (2008). Machine learning approach to automatic exudates detection in retinal images from diabetics' patients. *Journal of Modern Optics*. pp. 1-17.