Diabetic Retinopathy Classification using FP Growth Algorithm and Probabilistic Neural Network

P.Sureka[1], B.Pavithra[2,] L.Monisha[3,] G.Vimal Subramanian[4]
[1, 2, 3] *IV CSE, Department of Computer Science and Engineering,M.A.M. College of Engineering and Technology, Trichy*
[4] *AP, Department of Computer Science and Engineering,M.A.M. College of Engineering and Technology, Trichy-621 105*

Abstract

Diabetic retinopathy (DR) is an eye disease caused by the complication of diabetes and we should detect it early for effective treatment. As diabetes progresses, the vision of a patient may start to deteriorate and lead to diabetic retinopathy. As a result, two groups were identified, namely non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR). In this paper, to diagnose diabetic retinopathy, three models like Probabilistic Neural network (PNN), and frequent-pattern growth algorithm is described and their performances are compared. The amount of the disease spread in the retina can be identified by extracting the features of the retina(ROI extraction). The features like blood vessels, haemorrhages of NPDR image and exudates of PDR image are extracted from the raw images using the image processing techniques and fed to the classifier for classification.

Keywords - *Diabetic Retinopathy, Non-Proliferative Diabetic Retinopathy (NPDR), Proliferative Diabetic Retinopathy (PDR), Probabilistic Neural network (PNN), Frequent-Pattern Growth Algorithm (FP-Growth), ROI Extraction.*

I. INTRODUCTION

Diabetes is a group of metabolic diseases in which a person has high blood sugar, either because the body does not produce enough insulin, or because the cells do not respond to the insulin that is produced. [7] Diabetic retinopathy is one of the common complications of diabetes. It is a severe and widely spread eye disease. 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. Non-proliferative diabetic retinopathy is an early stage of diabetic retinopathy. In this stage, tiny blood vessels within the retina leak blood or fluid. The leaking fluid causes the retina to swell or to form deposits called exudates. Proliferative diabetic retinopathy, PDR is an attempt by the eye to grow or re-supply the retina with new blood vessels (neovascularisation), due to widespread closure of the retinal blood supply. [8] Unfortunately, the new, abnormal blood vessels do not re-supply the retina with normal blood flow, but bleed easily and are often accompanied by scar tissue that may wrinkle or detach the retina. The retinal image is taken in the RGB form by funds camera. A funds camera or retinal camera is a specialized low power microscope with an attached camera designed to photograph the interior surface of the eye, including the retina, optic disc, macula, and posterior pole.

Finally, Fuzzy C-means clustering is applied to segment the blood vessels in the image. After pre-processing of images is completed, features such as Radius, Diameter, Area, Arc length, Centre Angle and Half area are calculated for each image. Then Modelling Techniques like PNN, Bays Theory and SVM are used and their performances are compared. Finally, the images are classified into three groups namely, normal image, Non-Proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR).

II. RELATED WORK

During the recent years, there have been many studies on automatic diagnosis of diabetic retinopathy using several features and techniques. D. Vallabha et al. [1] proposed a method for automated detection and classification of vascular abnormalities in Diabetic Retinopathy using scale and orientation selective Gabor filter banks. R. Sivakumar et al. [2] presented a method to classify diabetic retinopathy subjects from changes in visual evoked potential spectral components. According to Thomas Walter et al. [3] exudates are found using their high grey level variation, and their contours are determined by means of morphological reconstruction techniques. HT Nguyenl et al. [4] proposed a multilayer feed forward network for the classification of DR. María García et al. [5] used a multilayer perceptron (MLP) classifier to obtain a final segmentation of Hard Exudates in the image. In [6], P. V. Nageswara road et al. proposed a new approach for protein classification based on a Probabilistic Neural Network and feature selection. S. Chaudhury et al. [7] address the problem of detecting

blood vessels in retinal images. They have used the concept of matched filter for detection of signals to detect piecewise linear segments of blood vessels in retinal images and constructed 12 different templates to search for vessel segments along all possible directions. In [8] Alireza Osareh et al. classified the segmented regions into two disjoint classes, exudates and non-exudates, comparing the performance of various classifiers. In [9] Wong Li Yun et al. Classified the four eye diseases using a three-layer feed forward neural network.

In [13], R. Priya and P. Aruna used SVM for the detection of diabetic retinopathy stages using colourfunds images. In [14] Computer-assisted automated red lesion detection was performed on digitized transparencies. [15] This paper investigates and proposes a set of optimally adjusted morphological operators to be used for exudates detection on diabetic retinopathy patients non-dilated pupil and low-contrast images. [16] described a method for automatically detecting new vessels on the optic disc using retinal photography. The 'matched filter response' method is a widely used template-based technique that uses a set of 2D Gaussian kernels with a fixed length and orientation to enhance the vessels. In [17], a post processing technique, based on edge detection, is applied to distinguish hard exudates from cotton wool spots and other artefacts. In [18] Keith A. Goatman et al. described a method for automatically deducting new vessels on the optic disc using retinal photography. Aliaa Abdel-Haleim et al. [23] presented a method to automatically detect the position of the OD in digital retinal funds images.

III. PREPROCESSING OF IMAGES

In detecting abnormalities associated with funds image, the images have to be pre-processed in order to correct the uneven illumination, not sufficient contrast between exudates and image background pixels and the presence of noise in the input funds image. The techniques for pre-processing include Gray scale Conversion, Adaptive Histogram Equalization, Discrete Wavelet Transform, Gaussian Matched Filter Response and Fuzzy C-means Clustering for segmentation of blood vessels.

A. Gray scale conversion

The acquired image resolution is 1280 × 1024 in 24bit JPEG format. The colour image of an eye is taken as input image and is converted to a greyscale image.

B. Adaptive histogram equalization

Adaptive histogram equalization which is used to improve contrast in images is applied to the gray scale converted eye image. Consider a running sub image W of $N \times N$ pixels cantered on a pixel $P(i, j)$, the image is filtered to produce another sub image P of $(N \times N)$ pixels according to

C. Discrete Wavelet Transform

The transform of a signal is just another form of representing the signal. It does not change the information content present in the signal. The Discrete Wavelet Transform (DWT), which is based on sub-band coding, is found to yield a fast computation of Wavelet Transform. It is easy to implement and reduces the computation time and resources required. Wavelet transform decomposes a signal into a set of basis functions.

D. The matched filter response

The matched filter is the optimal linear filter for maximizing the signal to noise ratio (SNR) in the presence of additive stochastic noise. The negative sign indicates that the vessels are darker than the background. Also, instead of 'n' different types of objects having to be identified, the problem reduces to decide whether or not a particular pixel belongs to a blood vessel. If the magnitude of the filtered output at a given pixel location exceeds a certain threshold, the pixel is labelled as a part of a vessel. Instead of matching a single intensity profile of the cross section of a vessel, a significant improvement can be achieved by matching a number of cross sections (of identical profile) along its length simultaneously.

F. The fuzzy c-means segmentation

Fuzzy C-means Segmentation (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. Here it is used to segment the input eye image and detect the blood vessels. Information about blood vessels can be used in grading disease severity or as part of the process of automated diagnosis of diseases with ocular manifestations.

G. Feature values

After performing the above mentioned pre-processing steps, the new eye image that is blood vessel detected image and Haemorrhages or exudates detected images are obtained. A set of feature values is taken from both blood vessel and Haemorrhages or exudates detected images. The feature values that are extracted are: Radius, Diameter, Area, Arc length, Center Angle and Half Area.

Features / Types of Eye Images	Radius (cm)	Diameter (cm)	Area (cm^2)	Arc length (cm)	Centre angle (ϕ)	Half area (cm^2)
Normal	144-156	260 - 320	65 - 80	5.7 - 7.6	2.2 - 2.6	32 - 39
NPDR	311 - 346	625 - 643	304 - 324	60.9 - 63.6	10.6 - 11.3	249 - 254
PDR	421 - 426	843 - 854	558 - 572	143 - 148	19.4 - 19.8	279 - 284

Table.1. The range of the feature values obtained for DR

IV. PROBABILISTIC NEURAL NETWORK

The PNN was first proposed in [11]. The PNN architecture is composed of many interconnected processing units or neurons organized in successive layers. [12] The input layer unit does not perform any computation and simply distributes the input to the neurons in the pattern layer.

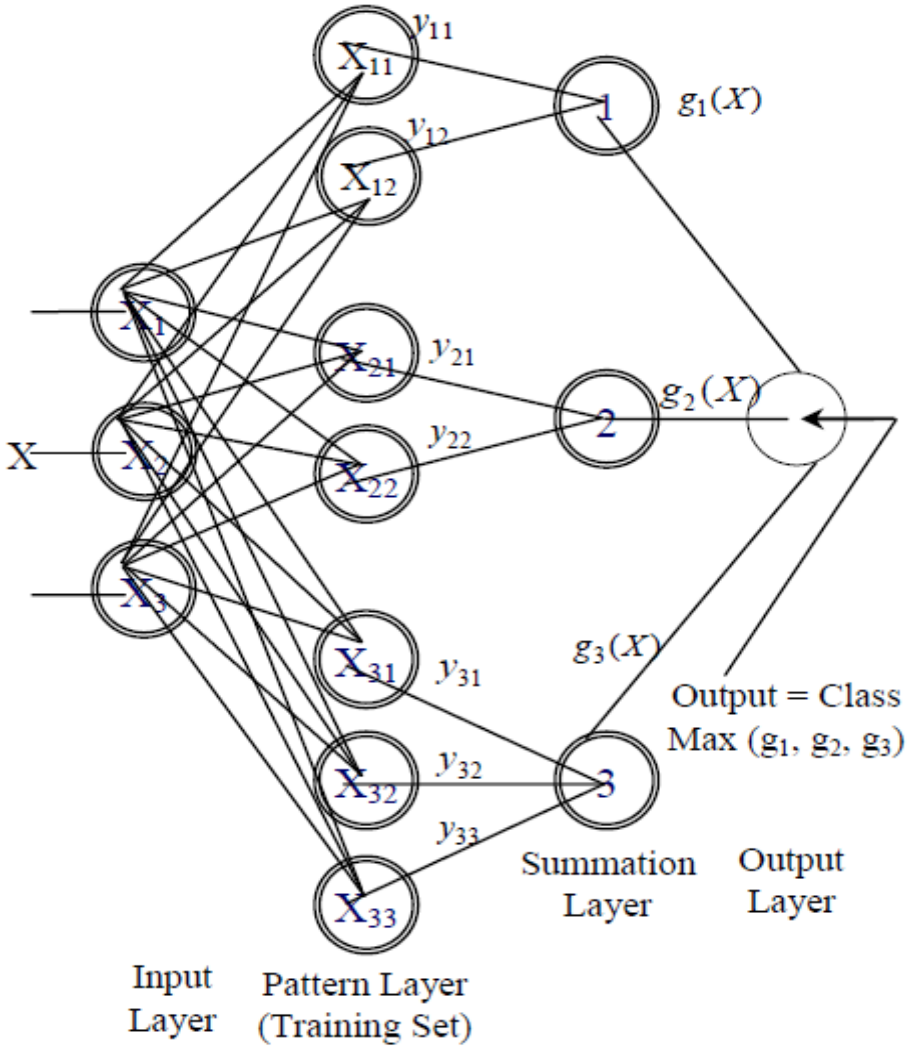

Figure 4.1 Architecture of Probabilistic Neural Network

Fig.4.1 shows a configuration of the PNN with four layers. There were six input features, which created a six dimensional input vector (X1, X2, X3, . . . , X6). Each image had a combination of specific values of the input vector—called an input pattern—that described the operating features of the image. The PNN model classifies that image from its input pattern into one of three categories (Normal, NPDR, PDR) as follows: In the input layer, the number of neurons is equal to the number of input features. In the pattern layer, the total number of neurons is the sum of the number of neurons used to represent the patterns for each category.

Each category may contain many training patterns (training vectors) whose dimension is equal to the number of input factors, and taking a set of specific values of input factors. The training vectors are imported from sample data and hence they are not always necessarily representative of all existing patterns for that class. However, this is the advantage of PNN, in that it can generalize to allow recognition of a new pattern of a class. The activation function used in the pattern layer, is the Gaussian kernel. In the summation layer, the number of neurons is equal to the number of categories. The outgoing signals can be adjusted according to loss and prior probability value. In the output layer, there is one neuron to represent the classification result. Suppose, Win is the input to the pattern layer for'*d*' varies from 1, 2,... 250, corresponding to 250 tested images and 'n' varies from 1, 2,....., 6 corresponding to the feature vector. The pattern layer can be processed and the output layer has a node for each pattern classification. The sum for each hidden node is sent to the output layer and the highest value wins. This method has been done for three classes namely normal, NPDR and PDR.

V. CONCLUSION

Diabetic Retinopathy is a disease which causes vision loss rapidly. To the input colour retinal images, pre-processing techniques like Greyscale conversion, Adaptive Histogram Equalization, Discrete Wavelet Transform, Matched filter Response and Fuzzy C-means segmentation are applied. After applying these pre-processing techniques the quality of the images are improved. From the pre-processed images, features were extracted for the classification process. As an achievement of this work, the DR has been classified into two categories NPDR and PDR using PNN, FP growth algorithm. All the three techniques used for the classification were good in performance, but SVM is more efficient than PNN and FP growth algorithm from the obtained results. Thus this work has given a successful Diabetic Retinopathy Diagnosing method which helps to diagnose the disease in early stage which mutually reduces the manual work. This infers that the FP growth algorithm model outperforms all other models. Also our system is run on 130 images available from "DIARETDB0: Evaluation Database and Methodology for Diabetic Retinopathy" and the results show that PNN has an accuracy of 87.69% and FP growth algorithm has an accuracy of 95.38%. However, we can improve the efficiency of the correct classification by extracting better features and by increasing the number of data in each class and also by combining with other pattern classification models.

REFERENCES

[1] D. Vallabha, R. Dorairaj, K. R. Namuduri and H. Thompson, "Automated Detection and Classification of Vascular Abnormalities in Diabetic Retinopathy", 38th Asilomar Conference on Signals, Systems and Computers, Vol. 2, pp. 1625 – 1629, 2004.

[2] R. Sivakumar, G. Ravindran, M. Muthayya, S. Lakshminarayanan and C. U. Velmurughendran, "Diabetic Retinopathy Classification", IEEE International Conference on Convergent Technologies for the Asia-Pacific Region, Vol. 1, pp. 205 - 208, 2003.

[3] T. Walter, J. C. Klein, P. Massin and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in colourfunds images of the human retina", IEEE Transactions on Medical. Imaging", Vol. 21, No. 10, pp. 1236 – 1243, 2002.

[4] H. T. Nguyen, M. Butler, A. Roychoudhry, A. G. Shannon, J. Flack and P. Mitchell, "Classification of diabetic retinopathy using neural networks", 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol. 4, pp. 1548 – 1549, 1996.

[5] MariaGarcia, Roberto Horner, Clara I. Sanchez, María I. Lopez and Ana Dies, "Feature Extraction and Selection for the Automatic Detection of Hard Exudates in Retinal Images", Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society Cite International, pp. 4969 – 4972, 2007.

[6] P. V. Nageswara Road, T. Uma Devi, D. S. V. G. K. Kaladhar, G. R. Sridhar, Allam Appa Road, "A Probabilistic Neural Network Approach for protein Super family Classification", Journal of Theoretical and Applied Information Technology, Vol. 6, No. 1, pp. 101, 2009.

[7] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson and M. Goldbaim, "Detection of blood vessels in retinal images using two-dimensional matched filters", IEEE Transactions on Medical Imaging, Vol. 8, No. 3, pp. 263 – 269,1989.

[8] Alireza Osareh, Majid Mirmehdi, Barry Thomas and Richard Markham, "Classification and Localisation of Diabetic-Related Eye Disease", Proceedings of the 7th European Conference on Computer Vision, pp. 502 – 516, 2002.

[9] Wong Li Yun, U. Rajendra Archery, Y. V. Venkatesh, Caroline Cheek, Lim Chook Min and E. Y. K. Ng "Identification of different stages of diabetic retinopathy using retinal optical images", Information Sciences, Vol. 178, No. 1, pp. 106 – 121, 2008.

[10] V. N. Vapnik, "Statistical Learning Theory", New York: Wiley, 1998.

[11] D. F. Specht, "Probabilistic neural networks", Neural Networks, Vol. 3, No. 1, pp. 109 – 118, 1990.

[12] K. Z. Mao, K. C. Tan and W. Ser, "Probabilistic Neural-Network Structure Determination for Pattern Classification", IEEE Transactions on Neural Networks, Vol. 11, No. 4, pp. 1009 – 1016, 2000.

[13] R. Priya and P. Aruna, "Automated Classification System for Early Detection of Diabetic Retinopathy in Funds Images", International Journal of Applied Engineering Research, Vol. 1, No. 3, 2010.

[14] Nicolai Larsen, Jannik Godt, Michael Grunkin, Henrik Lund-Andersen and Michael Larsen, "Automated Detection of Diabetic Retinopathy in a Funds Photographic Screening Population", Investigative Ophthalmology and Visual Science, Vol. 44, No. 2, pp. 767 – 771, 2003.

[15] Akara Sopharak, Bunyarit Uyyanonvara, Sarah Barman and Thomas H. Williamson, "Automatic detection of diabetic retinopathy exudates from non-dilatedretinal images using mathematical morphology methods", Computerized Medical Imaging and Graphics, Vol. 32, No. 8, pp. 720 –727, 2008.