

Detection of Diabetic Retinopathy using Image Processing and Machine Learning

Salman Sayed¹, Dr. Vandana Inamdar², Sangram Kapre³

M.Tech Student, Department of Computer Engineering, College of Engineering, Pune, Maharashtra, India¹

Head of Department, Department of Computer Engineering, College of Engineering, Pune, Maharashtra, India²

Research Analyst, Algo Analytics Financial Consultancy Pvt. Ltd., Pune, Maharashtra, India³

ABSTRACT: Diabetes is a chronic disease caused by either failure of insulin production in body or inability to resist insulin in the body. Complications of Diabetes leads to heart disorders, vascular disease and stroke, Kidney disease, Neuropathy and Diabetic eye disease known as “diabetic retinopathy”. Diabetes for a prolonged time damages the blood vessels of retina and thereby affecting seeing ability of a person and leading to diabetic retinopathy. Diabetic retinopathy is classified into two categories, non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR). In this paper, detection of diabetic retinopathy in fundus image is done by image processing and machine learning techniques. Probabilistic Neural Network (PNN) and Support vector machines (SVM) are the two models adopted for detection of diabetic retinopathy in fundus image and their results analyzed and compared. Accuracy of detection in SVM is 90% and that of PNN is 80%. Thereby SVM model is outperforming the PNN model.

KEYWORDS: Diabetic Retinopathy, Probabilistic Neural network, Support vector machine, Sensitivity, Specificity.

I. INTRODUCTION

Diabetes is a chronic disease which triggers the damage of retina. Diabetes is caused by lack of insulin in body or inability of body cells to respond to insulin in the body. Retinal blood vessels entering the retina from the optic disc are damaged which result in loss of sight. Initial stage shows no change in vision, but with time and lack of diabetes control may lead to total loss of vision. In this fast paced life diabetes is estimated to rise from 415 million in 2015 to 642 million in 2040. Diabetic retinopathy causes loss of sight in 1.8 million from 37 million people. Early stage of diabetic retinopathy known as Non proliferative diabetic retinopathy is shows signs of micro-aneurysms, exudates and hemorrhages. Micro-aneurysm are small red spots occurring in the retina due capillary swellings. Exudates are white or yellow deposits caused leaking fluid from the capillary. Hemorrhages are red spots occurring due to rupturing of capillaries and micro aneurysms. Later stage of diabetic retinopathy known as proliferative diabetic retinopathy is characterized by excessive growth of new blood vessels in the retina. Leakage of blood vessels lead to loss of vision.

II. RELATED WORK

As advancement has taken place, in recent years there has been an increased awareness and need for care and protection for eye vision. As such a lot of studies and research have taken place on automatic diagnosis of diabetic retinopathy using a number of advanced & more accurate precision techniques. To mention some of the important techniques developed are: D. Vallabha et al. [3] as per his studies has worked on a method on automated detection and classification of vascular abnormalities in Diabetic Retinopathy. The Method very well uses the scale and orientation selective Gabor filter banks for automated detection. DR. A. M. Mendonça, et al. [7] studies revealed a method in which microaneurysms can be validated. His method makes use of two new criteria, based on local intensity and contrast & shape relations. Huiqi Li and Opas Chutatape et al. [8] worked on the development of an automated fundus image processing system. The System helps highly to facilitate diagnosis by the ophthalmologists. Kenneth W. Tobin et al. [10] well developed a content-based image retrieval method. His method helps in verifying

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 1, January 2017

the hypothesis that retinal pathology can be identified and quantified. It makes use of visually similar retinal images from an available image archive. Chaudhury.S et. al [11] studies on the problems in detecting blood vessels in retinal images. In his Studies he has used the concept of matched filter for detection of signals to detect piecewise linear segments of blood vessels in retinal images. This lead to construction of 12 different templates for searching vessel segments along all possible directions. DR. V. Vijaya Kumari et al. [14] studies lead to the detection of many of the features such as blood vessel, exudates and optic disk and more importantly accurately using morphological operations. Yosawin Kangwanariyakul et al [17] studies revealed the usage of Back-propagation neural network (BPNN), the Bayesian neural network (BNN), the probabilistic neural network (PNN) and the support vector machine (SVM) all applied to develop classification models for identifying IHD patients. PV Nageswara rao et al [18] studied and advised a new approach for protein classification based on a Probabilistic Neural Network and feature selection. Lili Xu and Shuqian Luo et al [20] in their studies made use of a support vector machine (SVM) with Gaussian radial basis function as a classifier. As such they could identify hard exudates from digital retinal images. Priya.R and Aruna.P et al [21] used a support vector machine (SVM) for the detection of diabetic retinopathy stages. In this they used color fundus images.

III. PROPOSED SYSTEM

This paper presents an automated approach which speeds up the detection and classification of the disease diabetic retinopathy. A low power microscope named ophthalmoscope or the fundus camera is attached with a digital camera captures the image of interior surface of the eye which includes retina, optic discs, macula and the blood vessels. The images are captured by a Zeiss FF 450IR fundus camera. A modified digital back unit (Sony 3CCD color video camera) is connected to the fundus camera to convert the fundus image into a digital image. The images are usually obtained from the posterior pole's view including the optic disc and macula. Images had the resolution of 1280 x 1024 or 700 x 605 Pixels.

The proposed system classifies the image into categories such as normal, NPDR and PDR. For evaluation image set of 100 images were considered. The proposed system converts the RGB image obtained from the Fundus camera into grey image. Adaptive Histogram Equalization applied to the image improvises the contrast of the image. Followed by the adaptive histogram equalization, Noise in the image is removed by applying the matched filter response. The last step of the processing involves the clustering by Fuzzy c-means which makes the blood vessels of the image distinctly visible. After processing the image, features such as Radius, Diameter, Area, Arc length, Centre Angle and Half area are calculated for each image. Features extracted are passed on to the machines learning algorithms such as Support Vector machine and Probabilistic Neural Network for classifying the image into Normal, NDPR and PDR.

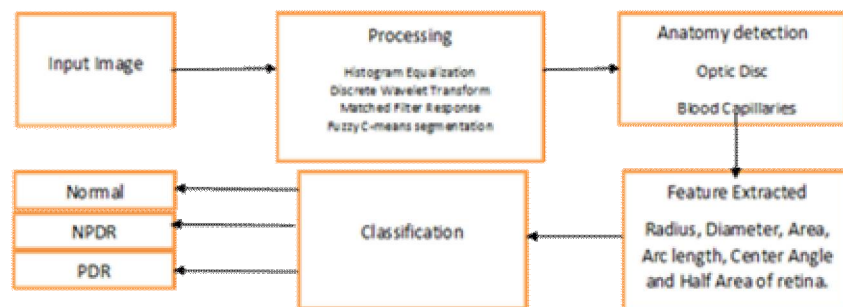


Fig. 1 Flow Diagram

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 1, January 2017

IV. PROCESSING

The Following are the Processing steps performed:

Image Acquisition: Image acquired from fundus camera is of 1280 x 1024 in 24-bit JPEG format.

Gray Scale Conversion: Image is converted into gray scale.

Adaptive Histogram Equalization: This computes several histogram and redistributes the lightness values across the image and thereby improves the contrast values of the image.

Discrete Wavelet Transform: Wavelets are discretely sampled and it captures both frequency and location information. Matched Filter Response: (MFR) reduces the noise in the image.

Fuzzy C-means Clustering: (FCM) technique is used to detect blood vessels from the sample fundus image. This method of clustering allows single fragment of data to be included in several clusters. Here it is used to segment the input eye image leading to detection of the blood vessels. Data obtained about blood vessels helps in grading the severity of the disease and automated detection of diabetic retinopathy.

Consider a fundus sub image 'W' of N x N of resolution P (i,j) which is filtered to produce sub image 'P' of (N x N).

$$P_n = 255[(O_w(p)-O_w(\text{Min})) / ((O_w(\text{Max})-O_w(\text{Min})))] \quad (1)$$

Here $O_w(P) = [1 + \exp((\mu_w - p) - \sigma_w)]^{-1}$

The Intensities given are the maximum and minimum in which local window mean is μ_w , and σ_w is standard deviation mean.

$$\mu_w = 1/N^2 * \sum P(i,j) \quad (2)$$

$$\sigma_w = \text{Sqrt} (1/N^2 * \sum (P(i,j) - \mu_w)^2) \quad (3)$$

The adaptive histogram equalization evenly brightens every part of the fundus image by making the dark part of the image bright and reducing the intensity of the brighter parts of the image.

The Discrete Wavelet Transform (DWT), is utilized for detection of blood vessels by performing fast computation on the basis of sub-band coding. This can be easily implemented on low end workstations with minimum computation time and resources. Wavelet transform breaks down the signal into collection of basis functions.

These basis functions are termed as wavelets. Mother wavelet $\psi(t)$ is utilized for dilations and shifting:

$$\Psi_{a,b}(t) = 1/\text{sqrt}(a) * \Psi((t-b)/a) \quad (4)$$

'a' being the scaling parameter and 'b' the shifting parameter.

The mother wavelet utilises the basic functions [cA, cH, cV, cD] = dwt2(X,'wname') which analyses and calculates the approximation coefficients matrix cA and here the coefficients matrices cH, cV, and cD (horizontal, vertical, and diagonal, respectively), achieved due to the wavelet decomposition of the input matrix X, where X represents the adaptive histogram equalized fundus image. Haar wavelet is utilized here which reduces the size of the image by half and the name given to the image is 'wname'. Thereby the size of the image now becomes 640 x 512. The Matched filter utilized here is optimal linear filter which reduces the noise and disturbances.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 1, January 2017

$$\text{hopt}(d) = -\exp(-d^2/2\sigma^2) \quad (5)$$

The negative sign indicated that the background is lighter than the blood vessels and helps in detecting whether the pixel belongs to the blood vessel or not. If pixel magnitude is greater than certain threshold, then it becomes part of the blood vessel.

MFR reduces the noise and disturbances the image.

Fuzzy c-means (FCM) is utilized for clustering which allows one fragment of data to two or more clusters. The image segment is analyzed whether it contains blood vessels or not. Blood vessels are then used to for detection of diabetic retinopathy

objective function used is:

$$J_m = \sum \sum u_{ijm} \|x_i - c_j\|^2, 1 < m < \alpha \quad (6)$$

Here m is any real number which is greater than 1, x_{ij} is membership of x_i to cluster k, c_j represents the center of the cluster, x_i is the 'i' th of d-dimensional measured data and objective function defines the similarity between data and cluster center.

The algorithm is composed of the following steps:

1. Initialize $U = [u_{ij}]$ matrix, $U(0)$
2. At k-step: calculate the centers vectors $C(k) = [c_j]$ with $U(k)$
 $C_j = (\sum u_{ijm} * x_i) / (\sum u_{ijm})$
3. Update $U(k)$, $U(k+1)$
 $u_{ij} = [1 / (\sum (\|x_i - c_j\| / (x_i - c_k))^{2/(m-1)})]$
4. If $\| U(k+1) - U(k) \| < \epsilon$ then STOP; otherwise return to step 2.

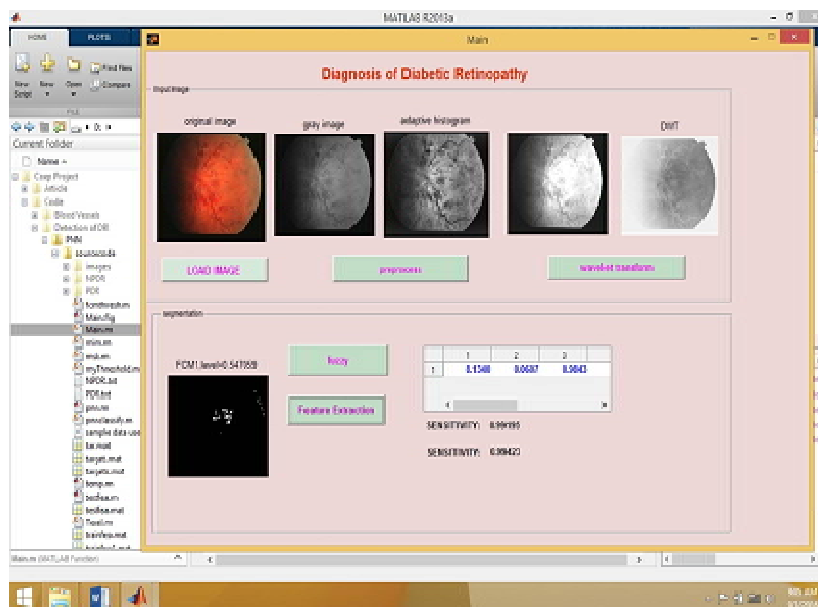


Fig. 2 Snapshot of System

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 1, January 2017

V. ANATOMY DETECTION - OPTIC DISC DETECTION USING HISTOGRAM MATCHING

For detection of optic disc accurately, we create a template for a optic disc. Three histogram, one for each color component is utilized for detection rather than using the image as a template. From the retinal image, the optic disc of size 80 x 80 pixels is extracted. Histogram for each color component is obtained for the extracted optic disc. 80 x 80 - pixel window is moved through the entire retinal image whose optic disc is to be founded. For every window the histogram for three color component is obtained. Correlation between the histogram of the channel in the window and the histogram of the template is calculated.

Correlation function between two histograms is expressed below:

$$c = 1 / (1 + \sum_i (a_i - b_i)^2) \quad (7)$$

Where a and b are two histograms that we want to calculate their correlation and c is the result of the correlation.

Therefore, if the two histograms (a and b) are similar

$$\sum_i (a_i - b_i)^2 \approx 0, \text{ and } c \approx 1, \text{ else and } c \ll 1.$$

VI. FEATURE EXTRACTION

Retinal image is circular in shape; hence circular features are extracted. Following the pre-processing steps, circular features of retinal images like Radius, Diameter, Area, Arc length, Center Angle and Half Area are calculated.

a) Radius: $r = (\text{Sqrt}(\text{Area} / \pi))$ (8)

b) Diameter: $d = 2 \pi r$ (9)

c) Area: $A = \pi * r * r$ (10)

d) Arc length: $\text{arclen} = R (2 \pi C) / 360$ (11)

Arc length = center angle of the arc
R = radius of the arc

e) Center Angle: $C = \text{Perimeter} / (360) (2 \pi r)$ (12)

f) Half Area: $\text{HF} = \text{Area of circle} / 2$ (13)

VII. SUPPORT VECTOR MACHINE

Support vector machine (SVM) is a supervised learning algorithm of machine learning. Machine or model is said to be learning if it utilizes the previous experience and the experience improves the performance of functions of machine or model. SVM is utilized to classify fundus images into Normal, NPDR and PDR classes.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 1, January 2017

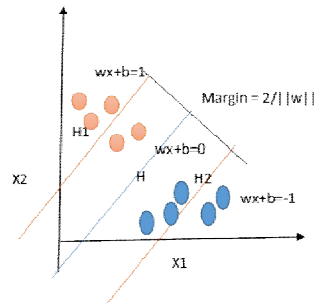


Fig. 3 Architecture of SVM

SVM is a supervised learning algorithm which breaks down the model into hyperplanes and classifies the objects. SVM model can be utilized for class 2, class 3 or more problems, in our case we have class 3 problem and are 3 classes are normal, NPDR and PDR. Training model is made for complete analysis of training data which tuned for extracting the features which can then be used for image classification. Features extracted from image are utilized by SVM for image classification. The image can be fragmented into various segments in terms of support vector classifiers. Kernel function is used to display data into hyperplane where hyperplane is disjunction between two classes. SVM can also perform non-linear classification by making use of non-linear kernel function. The non-linear kernel function maps data to higher dimensional space on which linear classifier can separate the input data.

Kernel function $K(x,y)$ gives inner product of $O(x),O(y)$ in feature space.

Polynomial kernel is given by: $K(x,x')=(xx'+1)^d$

Training vector utilized here is x and the kernel parameter is d . Input training vector is 100×6 . Classification is done into 3 categories namely normal, NPDR and PDR.

VIII. PROBABILISTIC NEURAL NETWORK

The Probabilistic Neural Network (PNN) is connective network of neurons which have the capacity to exchange data. The neurons are arranged in different layers. The different layers of PNN are input layer, pattern layer and output layer. The input layer retrieves the information passes it to pattern layer. The information passed from input layer to pattern layer is 'p' and the output from pattern layer is 'pij'

$$O_{ij}(p) = \frac{1}{(2\pi)^{d/2}} \frac{1}{q(d)} \exp \left[-\frac{(x-x_{ij})^T (x-x_{ij})}{2q^2} \right] \quad (14)$$

Here d is the pattern vector x dimension, and q is smoothing parameter, x_{ij} is neuron vector. For instance, if W_{dn} is input of tested images which ranges from 1 to 100 whereas 'n' varies from 1 to 6 corresponding to the feature vector. Pattern layer does the job of processing and output layer classifies the image into one of the three classes such as Normal, NPDR and PDR.

**International Journal of Innovative Research in Science,
Engineering and Technology**
(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 1, January 2017

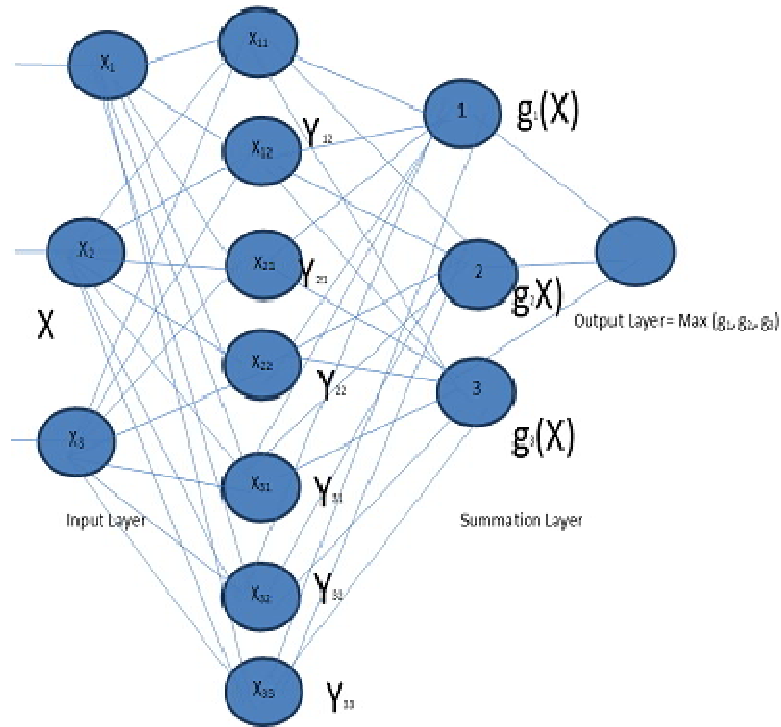


Fig. 4 Architecture of SVM

IX. RESULT

The Proposed method was implemented in Matlab R2013a. Sensitivity, Specificity and Percentage of accuracy for the three classes of eye images using the two classifier is found out.

Classes	No of Data Training	No. of data testing	No. Of correctly classified	Percentage Classification
Normal	10	10	10	100
NPDR	48	48	39	81.25
PDR	22	22	18	81.81
Average				87.68

Table 1: Result of SVM

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 1, January 2017

TN	TP	FN	FP	Sensitivity	Specificity
10	57	13	0	81.42%	100%

Table 2: Result of Sensitivity and Specificity

Classes	No of Data Training	No. of data testing	No. Of correctly classified	Percentage Classification
Normal	10	10	6	60
NPDR and PDR	20	15	12	80
Average				70

Table 3: Result of PNN

TN	TP	FN	FP	Sensitivity	Specificity
6	18	7	4	72%	71%

Table 4: Result of Sensitivity and Specificity

X. CONCLUSION

Diabetic Retinopathy is a disease which causes vision loss rapidly. The input color retinal images are of poor quality. So they were pre-processed using Grayscale conversion, Adaptive Histogram Equalization, Discrete Wavelet Transform, Matched filter and Fuzzy C-means segmentation. From the pre-processed images features were extracted for classification process. As an achievement of this work, the DR has been classified into two categories NPDR and PDR using PNN and SVM. The two techniques used for the classification were good in performance, but SVM is more efficient than PNN 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.

XI. FUTURE SCOPE

Augmented reality can aid in ophthalmologists in detection of diabetic retinopathy. Fundus images observed through screen of mobile phone or computer can show 3D augmented visuals of anatomical features of the retina. Regions where Ophthalmologists are not available, fundus images can be screened and augmented reality can aid the technician in generating reports of Diabetic retinopathy. We have prepared a prototype of retina scanner on android app which implements augmented reality features. When Eye anatomy picture observed through mobile 3D eye ball is augmented on screen showing eye anatomy features.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 1, January 2017

REFERENCES

- [1] Diabetes mellitus: http://en.wikipedia.org/wiki/Diabetes_mellitus
- [2] Advanced Vision Care: <http://www.advancedvisioncare.com/conditions.php>
- [3] Vallabha, D., Dorairaj, R., Namuduri K. and Thompson., H, "Automated Detection and Classification of Vascular Abnormalities in Diabetic Retinopathy", IEEE, 2004.
- [4] R. Sivakumar, G. Ravindran, M. Muthayya, S. Lakshminarayanan, C.U. Velmurughendran, "Diabetic Retinopathy Classification", IEEE, 2003.
- [5] Thomas Walter, Jean-Claude Klein, Pascale Massin, and Ali Erginay," A Contribution of Image Processing to the Diagnosis of Diabetic Retinopathy — Detection of Exudates in Color Fundus Images of the Human Retina", VOL. 21, NO. 10, IEEE Transactions on Medical Imaging, October, 2002,
- [6] HT Nguyen, M Butler, A Roychoudhry, AG Shannon, J Flack, P Mitchell, "Classification of Diabetic retinopathy using neural networks", 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Amsterdam 1996.
- [7] A.M. Mendonca, A.J. Campilho, J.M. Nunes, "Automatic Segmentation of Micro aneurysms in Retinal Angiograms of Diabetic Patients".
- [8] Huiqi Li and Opas Chutatape, "Fundus Image Features Extraction", Proceedings of the 22nd Annual EMBS International Conference, July 2328, 2000.
- [9] María García, Roberto Hornero, Clara I. Sanchez, María I. Lopez and Ana Díez, "Feature Extraction and Selection for the Automatic Detection of Exudates in Retinal Images", Proceedings of the 29th Annual International Conference of the IEEE EMBS Cite Internationale, Lyon, France, August 23-26, 2007.
- [10] Kenneth W. Tobin, Mohamed Abdelrahman, Edward Chaum, V. Priya Govindasamy, Thomas P. Karnowski, "A Probabilistic Framework for Content-Based Diagnosis of Retinal Disease", Proceedings of the 29th Annual International Conference of the IEEE EMBS Cite Internationale, Lyon, France August 23-26, 2007.
- [11] S. Chaudhuri, S. Chatterjee and N. Katz, "Detection of blood vessels in retinal images using two dimensional matched filters", IEEE Transactions on Medical Imaging, PP 263–269, 1989.
- [12] Jie Tian, Shanhua Xue, Haining Huang "Classification of Underwater Objects Based on Probabilistic Neural Network", Fifth International Conference on Natural Computation, IEEE, 2009.
- [13] Katia Estabridis and Rui J. P. de Figueiredo "Automatic Detection and diagnosis of diabetic retinopathy", IEEE, 2007.
- [14] V. Vijaya Kumari, N. Suriyanarayanan, C. Thanka Saranya, "Feature Extraction for Early Detection of Diabetic Retinopathy", International Conference on Recent Trends in Information, Telecommunication and Computing, 2010.
- [15] Alireza Osareh, Majid Mirmehdi, Barry Thomas, and Richard Markham, "Classification and Localisation of Diabetic-Related Eye Disease", Springer-Verlag Berlin Heidelberg, pp. 502– 516, 2002
- [16] Jian Wu, Guangming Zhang, Yanyan Cao, and Zhiming Cui, "Research on Cerebral Aneurysm Image Recognition Method Using Bayesian Classification", Proceedings of the 2009 International Symposium on Information Processing (ISIP 09), Huangshan, P. R. China, August 21-23, pp. 058-062, 2009.
- [17] Yosawin Kangwanariyakul, Chanin Nantasenammat, Tanawut Tantimongcolwat, Thanaokorn Naenna, "Data Mining of Magnetocardiograms for Prediction of Ischemic Heart Disease" EXCLI Journal, 2010.
- [18] Nageswara Rao Pv, Uma Devi T, Dsvdk Kaladhar, Gr Sridhar, Allam Appa Rao," A Probabilistic Neural Network Approach for protein Superfamily Classification", Journal of Theoretical and Applied Information Technology.
- [19] Inan Guler and Elif Derya U beyli, "Multiclass Support Vector Machines for EEG- Signals Classification", IEEE Transactions On Information Technology in Biomedicine, vol. 11, no. 2, March, 2007.
- [20] Lili Xu, Shuqian Luo," Support Vector Machine Based Method for Identifying Hard Exudates in Retinal Images", IEEE, 2009.
- [21] Priya.R, Aruna.P, "Automated Classification System For Early Detection Of Diabetic Retinopathy In Fundus Images", International Journal Of Applied Engineering Research, Dindigul, Volume 1, No 3,2010.
- [22] Vapnik. V, "Statistical learning theory Wiley", New York, 1998.
- [23] D. F. Specht, "Probabilistic neural networks," Neural Networks, vol. 3, no. 1, PP. 109–118, 1990.
- [24] C. Sinthanayothin, V. Kongbunkiat, S. Phoojaruenchanachai, A. Singalavanija," Automated screening system for diabetic retinopathy", in: 3rd International Symposium on Image and Signal Processing and Analysis, vol. 44, no. 2, 2003, pp. 767–771.