Diabetic Retinopathy Classification Using HARALICK Features

Reshma Sultana, K.S.Rajasekhar

Abstract: Diabetes is one of the most rapidly growing health threat around the world. Diabetic Retinopathy is abnormalities due to diabetes that affects eyes which leads to blindness over a period of time if not detected and cured in early stages. So detection and classification of Diabetic Retinopathy in early stages is important. In this paper, work is done on two database's one is DIARETDB0 database and other one is HRF (High-Resolution Fundus) database. The DIARETDB0 database consists of total 130 color fundus images, among which 20 are normal fundus images and 110 fundus images are having signs of Diabetic Retinopathy. The HRF (High-Resolution Fundus) database consists of total 45 color fundus images, among which 15 are normal fundus images, 15 fundus images are having signs of Diabetic Retinopathy and 15 fundus images are of glaucomatous patients. The texture features are extracted using Haralick Feature extraction Process. The Haralick Features are nothing but combination of Haar-DWT (Discrete Wavelet Transform) features and GLCM (Gray-level co-occurrence matrix) features that is original image is decomposed using Haar-DWT then sub-band images are produced. These sub-band images are used to extract features by GLCM. The proposed algorithm for feature extraction gives promising results. The Haralick Features are used to classify the normal and diabetic retinopathic images using a classifier. The performance of classification is calculated using the term Accuracy.

Keywords*: *Diabetic Retinopathy, Fundus images, Haralick Features, Classification, Accuracy.

I. INTRODUCTION

According to IDF (International Diabetes Federation), Diabetes is one of the most growing diseases around the world. It is estimated that 425 million people are living with diabetes all over the world. By 2025, the number of patients with Diabetes may rise to 629 million around the world. Among which 25.8 million people are from USA, 31.7 million people are from India and so on. The Indian Diabetic population may increase to 79.4 million by 2030 [1]. Since the people with diabetes are increasing day by day, the people with Diabetic Retinopathy are also increasing day by day, as Diabetic Retinopathy is abnormalities in retina due to diabetes that affects eyes, which even leads to blindness over a period of time. So detection in early stages is necessary to prevent blindness. Diabetic Retinopathy is caused due to damage of blood vessels of light-sensitive tissue at the back of retina. Over time, too much sugar in the blood can lead to blockage of tiny blood vessels that nourish retina, cutting off its blood supply. As a result of this, the eye attempts to grow new blood vessels. But these new blood vessels do not develop properly and leaks easily. There are two types of Diabetic Retinopathy, Early diabetic Retinopathy and Advanced Diabetic Retinopathy. Early Diabetic Retinopathy is most commonly called as non-proliferative diabetic retinopathy (NPDR). Due to NPDR, the walls of blood vessels in the retina get weaken. Tiny bulges called microaneurysms extend from vessel walls of smaller vessels which leads to leaking of fluid and blood in to retina and larger retinal vessels become irregular in diameter. As more blood vessels get blocked NPDR becomes severe. Advanced Diabetic Retinopathy is more severe type most commonly called as proliferative diabetic retinopathy (PDR). Due to PDR, damaged blood vessels close off because of which there will be growth of new abnormal blood vessels in retina from which there will be leakage of clear jelly like substance that fills center of eye. A scar tissue is accelerated due to growth of new blood vessels may cause retina to detach from back of eye. When these new blood vessels interfere with normal flow of fluid from eye the pressure may build in eyeball which damages the nerve that carries images to brain. Diabetic Retinopathy affects both the eyes. At first, Diabetic Retinopathy has no symptoms only there will be mild vision problems such as spots or dark strings floating in the vision, Blurred vision, Fluctuating vision, Impaired color vision, Dark or empty areas in the vision. But as time goes on it leads to complete vision loss [2].

Revised Manuscript Received on March 25, 2019.

Reshma Sultana, DEPT. of ECE, ANU college of Engineering and Technology, "Centre of Excellence in VLSI Design & Antennas", Acharaya Nagarjuna University, Guntur A.P, India.

K.S Rajasekhar, DEPT. of ECE, ANU college of Engineering and Technology, Acharaya Nagarjuna University, Guntur, A.P, India.

II. RELATED WORK

Ragav Venkatesan et al., (2012) [3], proposed a method for classification of diabetic retinopathy. In this the pre-processing is done by RGB conversion and K-means clustering. The feature extraction is done using auto correlogram and classification using MIL framework. The datasets used are DIARETDB0, DIARETDB1, STARE, and MESSIDOR. The accuracy achieved is 87% and is compared with other method accuracies. Kanika Verma et al., (2012) [4], proposed Detection and Classification of Diabetic Retinopathy. The detection was done using adaptive threshold, local threshold and Feature extraction was based on contrast enhancement. The classification of these extracted features were done by Neural Network. The dataset used is STARE dataset. The accuracy achieved is 88%. Mohith

Singh Solanki (2015) [5], proposed Feature extraction and classification using different methods. In Pre-processing, Image compression, adaptive histogram equalization and morphological operations are performed. Feature extraction is done using canny edge detection and Thresholding. These extracted features are given to deep neural network classifier. The dataset used is IDRiD (Indian diabetic retinopathy image dataset). The accuracy achieved is 55%.

R.S.Mangrulkar (2017) [6], proposed retinal image classification technique for diabetic retinopathy identification. In this method pre-processing is done using green channel image, edge detection is done using prewitt, sobel, canny methods, Kirsh process is used for segmentation. Feature extraction is done using SURF (speeded up robust features) algorithm. These extracted features are classified using Graph based Approach. The dataset used are DRIVE, INSPIRE-AVR, VICAVR. The accuracy achieved is 88%.

Salman Sayed (2017) [7], proposed diabetic retinopathy detection and classification using machine learning. In this method, pre-processing is done by RGB conversion and adaptive histogram equalization. Feature extraction is done using DWT and Fuzzy C-means clustering. These extracted features are used for classification using classifiers SVM and PNN, compared with each other. Images are acquired from fundus camera is of 1280x1024 in 24-bit JPEG format. The accuracies achieved are 87.68% and 70% respectively.

Ratul Ghosh et al., (2017) [8], proposed automatic detection and classification of diabetic retinopathy using CNN with three class classification. The dataset used is Kaggle dataset. The accuracy achieved is 85%.

Mohamed Chetoui (2018) [9], proposed Diabetic retinopathy detection. In this the feature extraction is done using LBP (local binary patterns), LTP (local ternary pattern), LESH (local energy based shape histogram) and compared with each other for accuracies. The classification is done to the extracted features of three using SVM. The dataset used is MESSIDOR. The accuracies are 80.5, 82.2 and 86.5 respectively.

Xiaoliang wang (2018) [10], proposed Diabetic Retinopathy detection and classification using CNN with a training algorithm of Inception Net V3. The dataset used is Kaggle Dataset. The accuracy obtained is 63.23%.

III. PROPOSED METHODOLOGY

Our Proposed methodology is shown in the Figure 1 in the form of a flowchart. This flowchart explains the total procedure of this paper step by step.

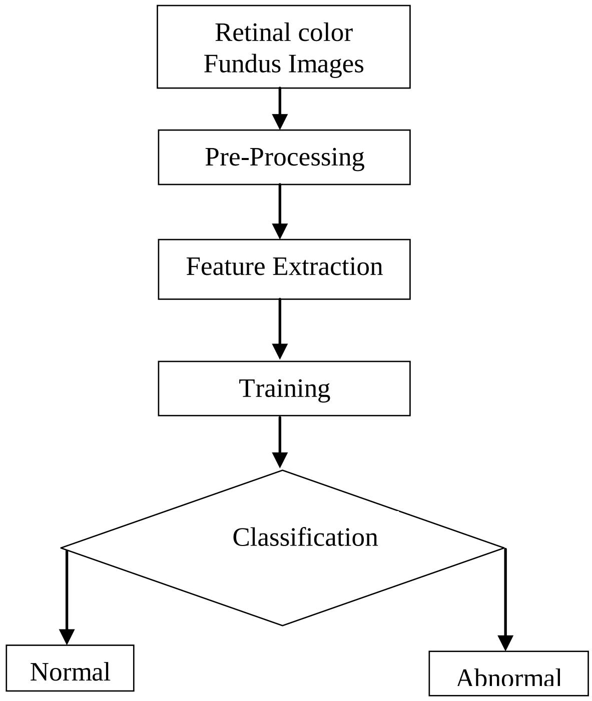

Figure 1: Block Diagram for proposed methodology

A. Image Database collection:

The images for the proposed methodology are collected from two databases. One is DIARETDB0 database, it consists of 130 color fundus images among which 110 images contains signs of diabetic retinopathy and 20 images are of normal. These images are in PNG (portable network graphics) format. For Present work 130 images are used for Classification. The other database is HRF (High resolution fundus) database, it consists of total 45 images, among which 15 images are normal, 15 images are of Diabetic Retinopathy and 15 images are of glaucomatous patients. These images are in JPG (joint photographic experts group) format. Figure 2 represents normal color fundus image of retina and Figure 3 represents Diabetic retinopathic image. The resultant accuracies of two databases are compared.

B. Pre-processing:

The Retinal images are either three dimensional or two dimensional images. So these three dimensional images need to be converted in to two dimensional format that means RGB to Grayscale format. The de-noising of the image can be done by passing images through median filter since the images are mostly affected by noise.

Figure 2: Normal image

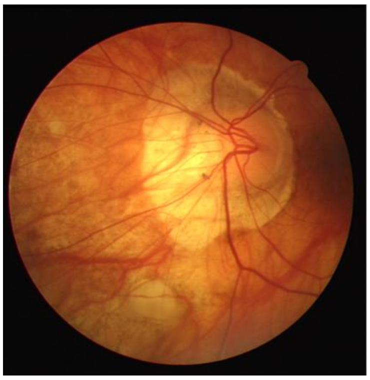

Figure 3: Abnormal image (Diabetic Retinopathy)

C. Feature Extraction:

By using feature extraction, the information such as color, shape, texture of an image can be obtained. Among these texture is a key component of human visual perception. The abnormalities in the retinal image are not visually distinguishable. So different Feature extraction algorithms are used for extracting the features of both normal and abnormal retinal images and based on that features, classification of normal and abnormal retinal images is done, so that diabetic retinopathy can be identified based on that classification. In this paper, the feature extraction algorithm used is Haralick (combination of Haar-DWT and GLCM). Haralick feature extraction is applied on retinal images of both DIARETDB0 database and HRF database, results of both databases are compared.

DWT (Discrete Wavelet Transform):

For many years, analysts used Fourier transform as a powerful tool, since it gives information regarding frequency content of signal. But this is not much effective for analysis because it doesn't give good frequency and time resolution at equal period. So wavelet transform is used since it gives good frequency and time resolution at equal period. There are different types of DWT's such as Haar-DWT, Daubechies-DWT, Symlets-DWT, Coeiflets-DWT, Biorthognal-DWT and Discrete Meyer Wavelet. Among these for the proposed method, Haar wavelet is used since Haralick features are combination of Haar wavelet and GLCM. The Haar wavelet is simplest one and is related to Haar transform which decomposes discrete signal in to sub-signals of half its length. The basic Haar Wavelet is represented as in Figure 4.

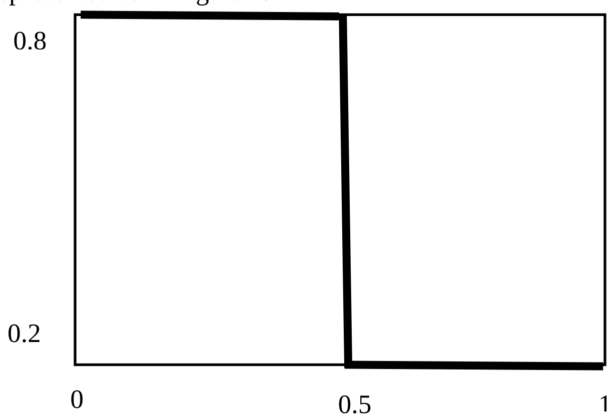

Figure 4: Haar Wavelet

The one-level DWT decomposition is represented as

$$\psi(p,q) = \psi(p)\psi(q) \quad (1)$$

$$\varphi^{H}(p,q) = \varphi(p)\psi(q) \quad (2)$$

$$\varphi^{V}(p,q) = \psi(p)\varphi(q) \quad (3)$$

$$\varphi^{D}(p,q) = \varphi(p)\varphi(q) \quad (4)$$

Where $\psi(p,q)$, $\varphi^{H}(p,q)$, $\varphi^{V}(p,q)$ and $\varphi^{D}(p,q)$ represents the approximated signal, signal with horizontal details, signal with vertical details, and signals with diagonal details respectively. The Two-level decomposition using DWT is represented in the form of flow diagram as shown in Figure 5.

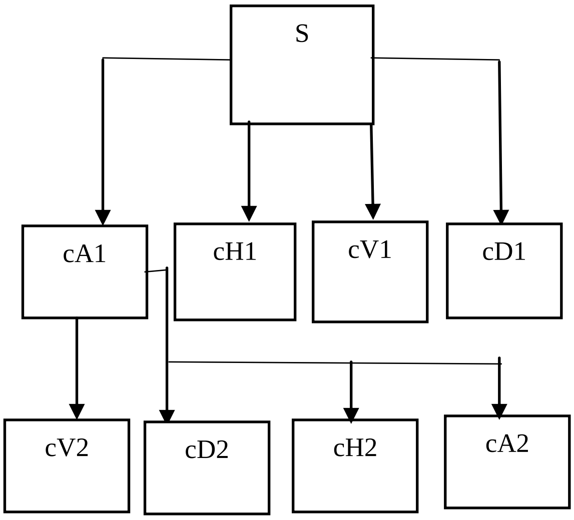

Figure 5: Two-level decomposition tree

GLCM (GRAY LEVEL CO-OCCURENCE MATRIX):

After using Haar wavelet on images only horizontal, vertical and diagonal decomposed approximations are considered. GLCM texture considers relation between two pixels at a time which are called as reference and neighbor pixel. The neighbor pixel is choosen to be one to the right of each reference pixel, this is also expressed as (1, 0) relation which means one pixel in X-direction and zero pixels in Y-direction. There are several textural features of GLCM which are defined using mathematical notations such as:

Contrast is defined as

$$\text{Contrast} = \sum_{p,q=0}^{M-1} R(p,q)(p-q)^2 \quad (5)$$

Correlation is defined as

$$\text{Correlation} = \sum_{p,q=0}^{M-1} R(p,q)\frac{(p-\mu)(q-\mu)}{\sigma^2} \quad (6)$$

Energy is defined as

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$\text{Energy} = \sum_{p,q=0}^{M-1} R(p,q)^2 \quad (7)$$

Homogeneity is defined as

$$\text{Homogeneity} = \sum_{p,q=0}^{M-1} \frac{R(p,q)}{1+(p-q)^2} \quad (8)$$

Entropy is defined as

$$\text{Entropy} = \sum_{p,q=0}^{M-1} -\ln(R(p,q))(R(p,q) \quad (9)$$

Variance is defined as

$$\text{Variance} = \sum_{p,q=0}^{M-1} R(p,q)(p-\mu)^2 \quad (10)$$

Inverse Difference Moment is defined as

$$\text{IDM} = \sum_{p,q=0}^{M-1} \frac{1}{1+(p-q)^2}\left(R(p,q)\right) \quad (11)$$

Mean is defined as

$$\text{Mean} = \sum_{p,q=0}^{M-1} pR(p,q) \quad (12)$$

Standard Deviation is defined as

$$\text{Standard Deviation} = \sqrt{\sum_{p,q=0}^{M-1} R(p,q)(p-\mu)^2} \quad (13)$$

Angular Second Moment is defined as

$$\text{ASM} = \sum_{p,q=0}^{M-1} \{R(p,q)\}^2 \quad (14)$$

Where $R(p,q)$ = Element p,q of normalized symmetrical GLCM.
M = Number of gray levels in image as specified by number of levels in under quantization on GLCM texture page of variable properties.

The above all the Textural Features are extracted as they provide more detailed information about medical images.

CLASSIFIER:

In our proposed paper, SVM (Support Vector Machine) is used for Classification. This Classifier is a binary classifier, it trains the data first and classifies the data in to two classes with hyperplane separation with complex boundaries depending on support vectors. The Classifier of SVM can be used in a modular manner (as the kernel function) and therefore depending on the purpose, domain and the separability of the feature space different learners are used. There are different kernels in the SVM such as Linear Kernel, Polynomial Kernel, RBF (Radial Basis Function) kernel and Quadratic Kernel. The Kernels are represented as follows:

Linear Kernel is defined as

$$G(\vec{r},\vec{s}) = \vec{r}\cdot\vec{s} \quad (15)$$

Where $\vec{r}$, $\vec{s}$ are feature vectors composed of binary word features.

Polynomial Kernel is defined as

$$G(\vec{r},\vec{s}) = (\vec{r}\cdot\vec{s}+b)^m, m>1 \quad (16)$$

RBF Kernel is defined as

$$G(\vec{r},\vec{s}) = \exp\left(-\frac{\left\|\vec{r}-\vec{s}\right\|^2}{2\sigma^2}\right) \quad (17)$$

Quadratic Kernel is defined as

$$G(\vec{r},\vec{s}) = (\vec{r}\cdot\vec{s}+b)^2 \quad (18)$$

The accuracies from different Kernels of SVM are compared.

IV. EXPERIMENTAL RESULTS

In this paper, two datasets are considered and the results of two datasets are compared with each other in terms of accuracy. First dataset consists of 130 images and second dataset consists of 45 images. Equal numbers of images are considered for training and testing. By using median filter noise is removed from dataset images then features are extracted using Haralick textural features such as contrast, correlation, energy, homogeneity, mean, standard deviation, Entropy, variance, IDM. All these features are imported in to MATLAB and then fed to classifier for training and testing. All the features are tabulated in Table I.

Table I: Features of Two databases

S.NO	FEATURES	DIARETDB0	HRF
1	Contrast	0.8531	1.2678
2	Correlation	0.8643	0.7736
3	Energy	0.8368	0.8369
4	Homogeneity	0.9763	0.9656
5	Mean	129.3653	167.7248
6	Standard Deviation	75.3413	79.7412
7	Entropy	0.4339	0.4141
8	Variance	2.0029e+03	1.2481e+03
9	IDM	2.4359e+05	6.7795e+05

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The accuracies at different kernels using two databases is compared in Table II and our proposed method is compared with existing methods which is shown in Table III.

Table II: Accuracies at different Kernels

Database	ACCURACY			
	Linear Kernel	RBF Kernel	Polynomial Kernel	Quadratic Kernel
DIARETDB0	84.615	80.769	81.0256	81.9231
HRF	86.363	77.272	71.2121	70.4545

Table III: Comparison of proposed method with existing methods

S NO	FEATURE EXTRACTION	CLASSIFICATION	ACCURACY
1	Canny edge detection	Deep neural networks	55
2	Fuzzy C-means	PNN	80
3	Haralick method	SVM	84(DIARETDB0)
			86(HRF)

V. CONCLUSION AND FUTURE SCOPE

In the Proposed method, the preprocessing is done by RGB to gray conversion and noise is removed by using median filter. The feature extraction is done using Haralick method and these extracted features are taken by classifier and based on these features classification is done using a supervised classifier. The accuracy of classification of normal and abnormal retinal images is calculated further in different kernels of SVM and are compared with each other. Further the results are compared with existing methodologies.

Feature Scope: To improve the accuracy by using different optimization techniques and classifiers such that better accuracy is achieved than proposed method.

REFERENCES

1. Statistics about number of diabetic patients https://www.statista.com
2. Diabetic Retinopathy types, causes and symptoms https://www.mayoclinic.org
3. Ragav Venkatesan, Parag Chandakkar, Baoxin Li, "Classification of Diabetic Retinopathy Images using Multi-class Multi-Instance Learning Based on color Correlogram Features", DOI: 101109/EMBC.202.6346216 ©2012 IEEE.
4. Kanika Verma, Prakash Deep and A.G. Ramakrishnan, "Detection and Classification of Diabetic Retinopathy using Retinal Images", © 2012 IEEE.
5. Mohit Singh Solanki, Dr. Amitabha," Diabetic Retinopathy Detection using Eye images", Indian Institute of Technology Kanpur, 2015.
6. R.S.Mangrulkar, " Retinal Image Classification Technique for Diabetes Identification", 2017 International Conference on Intelligent Computing and Control (I2C2).
7. Salman Sayed, Dr. Vandana Inamdar, Sangram Kapre, " Detection of Diabetic Retinopathy using Image Processing and Machine Learning", IJIRSET ,Vol.6, Issue 1, 2017.
8. Ratul Ghosh, Kuntal Ghosh, Sanjait Maitra, " Automatic Detection and Classification of Diabetic Retinopathy Stages using CNN", 4th International Conference on Signal Processing and Integrated Networks (SPIN),2017.
9. Mohamed Chetoui, Moulay A. Akhloufi, Mustapha Kardouchi, "Diabetic Retinopathy Detection Using Machine Learning and Texture Features", IEEE Canadian Conference on Electrical & Computer Engineering (CCECE), 2018.
10. Xiaoliang Wang, Yongjin Lu, Yujuan Wang, Wei-Bang Chen, "Diabetic Retinopathy Stage Classification using CNN", IEEE International Conference on Information Reuse and Integration for Data Science, © 2018 IEEE.

AUTHORS PROFILE

Author-1 Photo

Reshma Sultana is pursuing M. TECH with specialization in Communication Engineering and Signal Processing in ANUCET, Acharya Nagarjuna University. She completed her B. TECH from V R Siddhartha Engineering College in the year 2012-2016. Her interesting areas are Image Processing, Signal Processing and Communication.

Author-2 Photo

K.S.Rajasekhar is working currently as an Assistant Professor in the Department of ECE, ANUCET, Acharya Nagarjuna University. He completed his M. TECH from Vignan University, Guntur. He completed his B. TECH from Vijayanagar Engineering College, Karnataka. His Research area is Image Processing.

*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*