Indian J.Sci.Res. 17(2): 93-98, 2018 ISSN: 0976-2876 (Print)
ISSN: 2250-0138(Online)

PCA AND ICA MODEL FOR IMPROVING RETINAL IMAGE CLASSIFICATION IN DIABETIC RETINOPATHY

[1]M. Navinkumar, [2]T. Swetha, [3]S. Karthikeyan
[1,3]DepartmentofElectronics& Communication Engg,Sri Krishna College of Tech., Coimbatore
[2]Department of Electronics &InstrumentationEngg, Keshav Memorial Institute of Tech., Hyderabad

Abstract-Screening of Diabetic Retinopathy (DR) with proper time treatment prevents blindness. Various researchers have focused on their work on the enhancement of computer-aided lesion-specific detectors. DR generates various types of retinal lesions. The diagnosis of DR depends on the detection of retinal lesions such as microaneurysms and exudates in retinal images obtained by a fundus camera. However, bright lesions such as exudates can share similar appearances while being signs of several diseases. However, DR does not exhibit any distinctive symptoms in which the patient can easily observe until a severe stage is reached. It is, therefore, important to develop some advanced DR detection technologies. Some of the detection methods are designed by using this research methodology. The experimental results with 51 number of real-time fundus image datasets, which was collected from the "Eye Foundation" hospital using ZEISS FF450 plus camera. The proposed approach delivers excellent discriminative results for normal and exudates images. It greatly achieves 97.3510% of accuracy, 94% of sensitivity, and 99.0099% of specificity for hemorrhages images.

Keywords— PCA, ICA, Mahalanobis distance, PLS

I. Introduction

The number of blind people is increasing rapidly at higher rates than ever before. Recent studies have proven that the number of cases with diabetic retinopathy (DR) has increased by 89%. To provide healthcare for the high number of diabetics, a means for large scale retinal screening systems will be needed. This work presents the use of Independent Component Analysis to distinguish between the normal and infected retinal images.

Various image processing techniques are being used to detect the Diabetic Retinopathy. These methods help the doctors for offering better treatment for the patients and early assessment of diabetic retinopathy. But still, in most of the techniques, the identification of DR is not accurate.The existing multi-scale amplitude modulation and frequency modulation (AM-FM) methods have less accuracy while identifying the normal and the pathological retinal images. This technique is implied to detect the microaneurysms, exudates, neovascularization on the retina, hemorrhages, normal retinal background and normal vessel patterns on theretina. In this section, Optic Nerve Head (ONH) is detected based on the edge information. However, while detecting the ONH, masking is not performed. It verifies the features of disease by using a pre-defined value from AM and FM based feature extraction method.The results obtained in this technique have a statistical differentiation of normal retinal structures shows significant capability for use in automatic DR screening. In this work, a novel method that employs the AM-FM decomposition along with Independent Component Analysis is used to normalize most of the approximations.

Diabetes causes an excessive amount of glucose to remain in the blood stream which may cause damage to the blood vessels. Diabetes also increases the risk of having glaucoma, cataracts, and other eye problems. The World Health Organization (2015) calculated that, currently 135 million people have diabetes mellitus worldwide and the number will increase to 300 million by the year 2025. A side effect of diabetes is Diabetic Retinopathy in which different parts of the retina get affected. These methods help the doctors for offering better treatment for the patients and early assessment of diabetic retinopathy.

[1]Corresponding Author

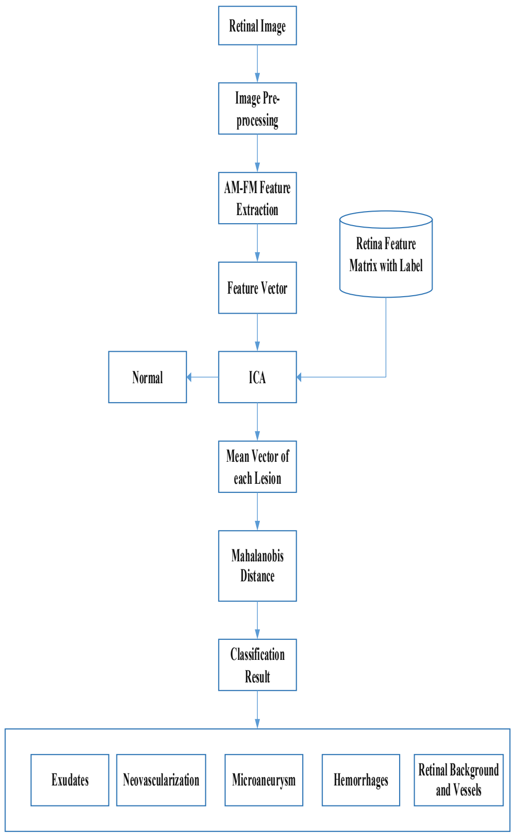

Figure 1:Overall Flow Diagram for ICA and UDA for DR Lesion detection

Initially, the retinal images are collected and loaded to the proposed ICA-module. The existence of noise and size variations in the images affect the further processing. Prior to detection of abnormalities and the features, the noise and background removal play the major role to improve the quality of the image with in order to achieve the better classification accuracy. The main intent of preprocessing is to normalize the captured retinal image for attenuating the image variations. Then, the image is approximated with the modulation components by using Amplitude Modulation (AM) and Frequency Modulation (FM) decomposition. The utilization of Instantaneous Frequency (IF) and Instantaneous Amplitude (IA) measures extract the features for the images.

The derivatives from the dominant neighborhood angle called relative angles to measure the geometry of texture features. The extraction of AM-FM components is scaled in different frequency ranges (filter banks). The different filters utilization consider the orientation of features that are encoded. Then, the cumulative distribution functions of IA and IF are used to characterize the retinal structures. Once the feature extraction is over, the Independent Component Analysis (ICA) is applied to reduce the dimension and find the specific lesion. The mean of the collection of regions plays the major role in lesion prediction and it depends on the Mahalanobis distance measurement. Finally, the Pixel Level Snake (PLS) is employed for classification of images such as MAs, hemorrhages, exudates, neovascularization (NV), retinal background and vessels. The classification result is further enhanced by application of uncorrelated discriminant vectors extraction based on fisher vector formulation in UDA process. The hybrid utilization of ICA-UDA in proposed work improves the classification accuracy significantly. This chapter discusses the workflow of each module in detail.

II.Related Work

In modern living style, lot of people are getting affected by diabetes. The diabetes is a systematic and chronic end organ disease, which occurs when the pancreas does not secrete enough insulin or the body is unable to process it properly. Diabetes causes an excessive amount of glucose to remain in the blood stream which may cause damage to the blood vessels. Diabetes also increases the risk of having glaucoma, cataracts, and other eye problems. The World Health Organization (2015) calculated that, currently 135 million people have diabetes mellitus worldwide and the number will increase to 300 million by the year 2025. A side effect of diabetes is Diabetic Retinopathy in which different parts of the retina get affected.

Diabetic retinopathy (DR) is a common eye disease, which occurs in most of all diabetic cases. It affects most of the people those who have diabetes for 10 or more years. The number of people being affected by this disease continues to grow in an alarming rate. DR is a silent disease, because it may only be recognized by the patient when the changes in the retina have progressed to a level where treatment is complicated and nearly impossible [7]. The prevalence of retinopathy varies according to the age of onset of diabetes and the duration of the disease. DR causes a fluid to leak into the macula region of the retina, which causes the retina to swell and leads to blur vision. Diabetic Macular Edema (DME) is an advanced symptom of diabetic retinopathy, which causes permanent vision loss. Identification of lesion in the early stage is an effective step for examining the disease. Many treatments are available to detect the DR in the initial stage. The most common retinal lesions are Micro aneurysms, Hemorrhages, Exudates and Neovascularization.

Micro aneurysms are small swellings that forms in the walls of tiny blood vessels on the retinal surface. Those small swellings may break which causes blood to leak into the nearby tissues. Hemorrhages are situated in the middle layer of the retina. A disorder, in which bleeding occurs into the light sensitive tissue on the back wall of the eye is called as retinal hemorrhages. Exudates appear as yellow or white structures in the retina. Based on the appearance and occurrence, there are two types of exudates such as hard exudates and soft exudates. Hard exudates have well defined boundaries. The soft exudates, also known as cotton wool spots contains unclear boundaries [9]. Neovascularization causes new blood vessels to grow due to extensive lack of oxygen in retinal capillaries.

Proliferative retinopathy is an abnormal growth of new blood vessels and it can be seen in the eye as neovascularization. The neovascularization is quite dangerous because the vessel grows abnormally out of the retina into the clear vitreous gel. The vessels grow beyond the supporting structure of retina and are very prone to bleeding, when they occur near the disc [4]. [5] proposed the Fuzzy C-Means (FCM) method for detecting the diabetic retinopathy edema. The input images were pre-processed using the median filter, hue saturation value and FCM. The median filter was used for minimizing the noises in the image. The median filter validated every pixel with its neighboring pixel for confirming that it was the representative of the surrounding.[10] proposed a median filtering framework for reducing impulse noise from gray scale images. The Standard Median Filter (SMF) minimized the corruption level and also modified the corrupted pixel intensity.[3] proposed a novel median filter for extracting the natural pixels for restoration. The suggested filter restored the images that had 1-99% of salt-and-pepper impulse noise.[11] suggested a median filter based de-noising algorithm for the color and grayscale images. The suggested median filter based algorithm removed the noisy pixels by the median value. The issues related to increasing the window size was removed.[2] suggested an applied median filtering technique for removing the noises in the MRI, cancer, X-ray and brain images. Based on the type of the noise and type of the filtering technique, the noises in the medical images were removed.

II.I.Frequency Scale And Filter Banks

The AM-FM components are extracted from different image scales. Consider the usage of 25 band pass filters associated with four frequency scales and nine possible Combination of Scales (CoS). Estimate AM-FM component over each combination of scales using Dominant Principle Component Analysis (DPCA). At lower frequency scales, the magnitude values of |IF| are small, and the extracted AM-FM features slowly reflect the variation of image textures.

Table 1:Band pass Filter associated with Multiple Image Scales

Frequency Scale Band	**Filters**	**Instantaneous Wavelength (period) Range in Pixels**	**Range in mm**
Low Pass Filter (LPF)	1	22.6 to ∞	0.226 to ∞
Very Low Frequencies (VL)	20-25	11.3 to 32	0.113 to 0.32
Low Frequencies (L)	14-19	5.7 to 16	0.057 to 0.16
Medium Frequencies (M)	8-13	2.8 to 8	0.028 to 0.08
High Frequencies (H)	2-7	1.4 to 4	0.014 to 0.04

The usage of different scales also considers size variability among structures such as MAs, exudates, hemorrhages, etc. For Diabetic Retinopathy (DR) patients, the lesion size varies. The dark lesions that are termed as Mas, and the light lesions are termed as exudates. The various filters that are being used consider the orientation of the features that are being encoded. Table 1 shows the frequency ranges of the band pass filters.

II.II.EncodingThe Structures Using Am-Fm

The retinal structures are used by introducing the Cumulative Distribution Functions (CDFs) of the IA, |IF| and the relative angle. Subsequently, the range of CDF values for each retina is estimated and varied according to the utilization of CoS. Based on the global minimum value and the global maximum value, the histograms are calculated. Low IA values are used to characterize the region with small pixel intensity variation. This characterization is done based on a higher frequency scales. It happens due to the presence of low-intensity variation region in weak frequency components.

The darker region in the retina has low IA values, which is present in the lower frequency scales. By using this concept, the exudates that are present with MAs can be easily identified. For instance, the background of the retina is analyzed in the entire frequency spectrum, which has rough constant low IA values. Generally, for a given frequency scale, the low IA values reflects the frequency components, which are present in the specific scales. Defining Retinal Characteristics of AM-FM Feature Vectors

It explains the estimation of AM-FM encode structures and the relation of encoding to the formation of related feature vectors for the detection of the analyzed lesions. The Instant Frequency magnitude (|IF|) is indifferent to the direction of image intensity differences. Moreover, the IF magnitude can be defined as a function of local geometry, which is dissimilar to the gradual-varying brightness variations. Hence, a single dark round structure in a lighter background has the same |IF| distribution. In the darker region, the |IF| distribution has a same single bright round structure of equal size.

This above condition is incorporated for exudates that are bright lesions and microaneurysms i.e. dark lesions. |IF| assessments are used distinguishing the two regions where one contains a single vessel similar to a normal retinal vessel. Another one has a various narrow vessel same as the vessels present in neovascularization. The information present in this two regions has the same frequency ranges. The second region has a high range of histogram when compared to the former region. This high count states the point that a larger number of pixels that reveal these frequency components are present in neovascularization.

II.III.ClassificationOf Retinal Images Using Ica&Uda

Figure 3 shows the steps involved in the classification of the retinal images. Initially, the features are extracted using AM-FM. The reduction of dimensionality is done using ICA. Further, the dimensionality is reduced using k-means clustering. Finally, Pixel Level Snakes (PLS) is applied to classify the images.

Independent Component Analysis (ICA)

In ICA, therandomly observed vector are indicated as $X = [X_1, X_2, \dots . X_m]^T$ whose elements are the combinations ofm independent elements of a random vector$[S_1, S_2, \dots . S_m]^T$, which is given by

$$X = AS \qquad (3)$$

Where A represents a $m \times m$ mixing matrix, the example values of X_j is indicated by $x_j$and$j = 1,2, \dots , m$. The aim of ICA is to detect the unmixing matrix$W$, which is the inverse of$A$. The best possible approximation of S is estimated as follows,

$$Y = WX \cong S \qquad (4)$$

In this research work, five assumptions are made to use ICA. The initial assumption is statistical independence that is applied between each of the sources S_i from the sources vector S. Next, the number of mixtures must be equal to the number of sources. These mixtures should be independent of each other. The mixing matrix should be square and full rank. The third assumption is, the model should be free from noise, and it is an only source of stochasticity in the model in the source vector S. In the fourth assumption, data are centered, which is also called as zero mean. In this assumption, the data must be preprocessed before it is used, and the observation vector is also whitened. Final assumption, the source signals should not have a Gaussian probability density function (pdf) except for the source of Gaussian.

2.3.1.Independence

A key concept that constitutes the foundation of independent component analysis is statistical independence. To simplify the above discussion consider the case of two different random variables s1 and s2. The random variable s1 is independent of s2 if the information about the value of s1 does not provide any information about the value of s2 and vice versa. Here s1 and s2 could be random signals originating from two different physical processes.

2.3.2. Non-Gaussianity and Independence

According to the central limit theorem, the distribution of a sum of independent signals with arbitrary distributions tends toward a Gaussian distribution under certain conditions. The sum of two independent signals usually has a distribution that is closer to Gaussian than the distribution of the two original signals. Thus, a Gaussian signal can be considered as a linear combination of many independent signals.

II.IV.Alternate Solution Using Uda (Uncorrelated Discriminant Analysis)

The UDA replaces the above proposed solution. The change is made in the reduction of dimensionality of the original image after being encoded using AM-FM reduction.

2.4.1 Fisher Vector

Let $X = \{x_t, t = 1, \dots T\}$ be the set of T local descriptors extracted from an image. Assume that the generation process of X can be modeled by a pdfμ_λ with parameters λ^4. Gcan be described as a gradient vector

$$G_\lambda^X = \frac{1}{T} \nabla_\lambda log \, \mu_\lambda(X) \qquad (10)$$

The gradient of the log-likelihood describes the contribution of the parameters x_t the generation process. The dimensionality of this vector depends only on the number of parameters in λ, not on the number of patches$T$.

$$K(X, Y) = G_\lambda^{X\prime} F_\lambda^{-1} G_\lambda^Y \qquad (11)$$

Where F_λ is the Fisher information matrix ofμ_λ

$$F_{\lambda} = E_{x \sim \mu_{\lambda}}[\nabla_{\lambda} log\, \mu_{\lambda}(x) \nabla_{\lambda} log\, \mu_{\lambda}(x)'] \quad (12)$$

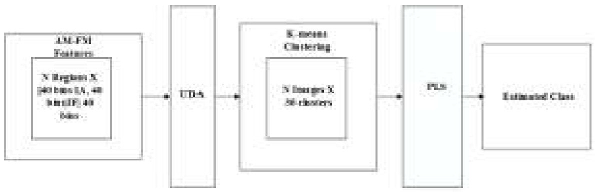

Figure 2:Classification of Retinal Images

2.4.2 Uncorrelated Discriminant Vectors

Let φ_1 be the Fisher vector. Suppose j directions $\varphi_1, \varphi_{2,\ldots\ldots}\varphi_r(r \geq 1)$ are obtained. To attain the uncorrelated discriminant features, we can let the $(r+1)^{th}$ direction φ_{r+1} which maximizes the Fisher criterion function $F(\varphi)$ with the following conjugated orthogonality constraints

$$\varphi_r^T +_1 S_t \varphi_i = 0 \text{(i=1,2....r)} \quad (13)$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_k \end{bmatrix} = \begin{bmatrix} \varphi_1^T \\ \varphi_2^T \\ . \\ . \\ . \\ \varphi_k^T \end{bmatrix} X \quad (14)$$

The above transformation along with the constraint is also known as uncorrelated discriminant transformation.

Uncorrelated Feature Extraction

The main goal of feature extraction is to choose all microaneurysms that exist in the pre-processed image. Microaneurysms look like isolated patterns, which are separated from the vessels. The structures of microaneurysms can be mined based on shape, size and intensity level. Microaneurysms are dark reddish in color and look as small red dots, which has 10 to 100 microns diameter and are circular in shape.

Once the image is pre-processed, by discriminating the pre-processed image from the blood vessels, the identified microaneurysms are segmented. Usually, MA and vessels both appear in a reddish color, and the only difference is MAs cannot exist on vessels. Blood vessels in the retina are larger in the area and are linked component, the intensity is related to MA. Finally, the threshold value is finalized by investigation process. The higher threshold is eliminated by computing the respective objects to remove the unwanted blood vessels.

2.4.3 Classification of Retinal Images using UDA

Figure 3 shows the steps involved in the classification of the retinal images. Initially, the features are extracted using AM-FM. The reduction of dimensionality is done using UDA. Further, the dimensionality is reduced using k-means clustering. Finally, the Pixel Level Snakes (PLS) is applied to classify the images.

II.V.ProcedureTo Find Mahalanbolis Distance

Initially, the features for a particular region are extracted. Then, the dimensionality is reduced using Independent Component Analysis (ICA). It is performed after the mean of the collected regions corresponding to a specific lesion is found. The ICA reduces the noise as well the dimensionality of the image. Initially, we consider a statistical model

$$Y_y = A.x \quad (15)$$

Where x is considered as a random vector from m independent component analysis. $A$is an $m \times m$ matrix of invertible parameters. An observed vector with m components is represented as Y_y. Based on a set of *N* independent component, the parameter matrix A is estimated. From the estimation of A, the values corresponding to x and y by solving a linear system of equations. The distribution of x is unknown and the maximum likelihood of A need to be estimated.

Consider the population version of ICA in which $p^*(y)$ denotes the true distribution of y and $p(y)$ denotes the model. The Kullback-Leibler (KL) divergence between the distributions is minimized as $D(p^*(y)||p(y))$. Define $W = A^{-1}$, so that x=Wy. Since the KL divergence is invariant with respect to an invertible transformation, the problem is set equivalent to minimization of $D(p^*(x)||p(x))$. Let $\tilde{p}(x)$ denote the joint probability distribution obtained by taking the product of the marginal of $p^*(x)$.

$$D(p^*(x)||p(x)) = D(p^*(x)||\tilde{p}(x)|| \, p(x)),$$

for any distribution p(x) with independent components. Consequently, for a given A, the minimum over all possible p(x) is attained precisely at $p(x) = \tilde{p}(x)$. The minimal value is $D(p^*(x)||p(x))$ is exactly the mutual information between the components of x=Wy. In such a way, six mean values were found. From the six means Mahalanobis distance is estimated.

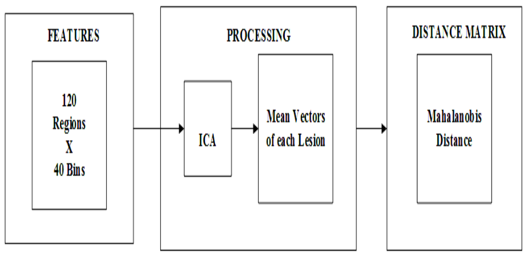

Figure 3:FindingMahalanobis Distance

III.Experimental Results

The experimentation shows the impact of the ICA over the AM-FM feature extraction.For the purpose of analysis, the ETDRS database is used. The proposed system uses ICA, which experiments against the existing systems. These existing systems use the PCA to provide results. Also, it involves the analysis of a system using UDA also. The analysis is done based on the level of accuracy based on Mahalanobis Distance and Hellinger Distance. The performance analysis section depends on the four parameters such as dimensionality (PCA), projection vector setup (ICA)) the number of components involved in mixture models and the confidence level of SVM classifier. The eigenvalues estimation in PCA play the major role in dimensionality in PCA. The components are retained in PCA with the eigenvalue is greater than 1. Table 2 illustrates the parameters required for the validation of proposed research work. Table 2 shows the result of image 1 and image 2 using PCA, ICA,. By considering image 1, the PCA approach does not detect that the given original input image 1 is DR or not in an efficient manner. Because, it detects mismatched result in our dataset. In some cases, the ICA technique is insufficient to detect the DR. The original input image 2 is taken into account to justify this statement.

References

[1] AbdulazizSalehYeslem Bin-Habtoor SSA-a 2016, ‘Removal Speckle Noise from Medical Image Using Image Processing Techniques’, International Journal of Computer Science and Information Technologies (IJCSIT), vol. 7,no. 1, pp. 375-377.

[2] BhausahebShinde DM, A.R. Dani 2012, ‘Study of Noise Detection and Noise Removal Techniques in Medical Images’, .J. Image, Graphics and Signal Processing, vol. 2, pp. 51-60.

[3] Hsieh M-H, Cheng F-C, Shie M-C &Ruan S-J 2013, ‘Fast and efficient median filter for removing 1–99% levels of salt-and-pepper noise in images’, Engineering Applications of Artificial Intelligence, vol. 26,no. 4, pp. 1333-1338.

[4] John Walsh RR 2015, About diabetics. Available from:<http://www.diabetesnet.com/about-diabetes/diabetes-complications/eye-complications>.

[5] Mohan A &Moorthy K 2013, ‘Early detection of diabetic retinopathy edema using FCM’, International Journal of Science and Research (IJSR), India, vol. 2,no. 5, pp. 115-118.

[6] Morales Martinez S, NaranjoOrnedo V, ANGULO LOPEZ J &Alcañiz Raya ML 2013, ‘Automatic detection of optic disc based on PCA and mathematical morphology’, IEEE transactions on medical imaging, vol. 32,no. 4, pp. 786-796.

[7] Qi X, Shen L, Zhao G, Li Q &Pietikäinen M 2015, ‘Globally rotation invariant multi-scale co-occurrence local binary pattern’, Image and Vision Computing, vol. 43, pp. 16-26.

[8] Shrestha S 2014, ‘Image Denoising using new Adaptive based [8]Median filter’, Signal & Image Processing : An International Journal (SIPIJ), vol. 5,no. 4, pp. 1-13.

[9] Sim DA, Keane PA, Tufail A, Egan CA, Aiello LP & Silva PS 2015, ‘Automated Retinal Image Analysis for Diabetic Retinopathy in Telemedicine’, Current diabetes reports, vol. 15,no. 3, pp. 1-9.

[10] Teoh SH & Ibrahim H 2012, ‘Median filtering frameworks for reducing impulse noise from grayscale digital images: a literature survey’, Int. J. Future Comput. Commun, vol. 1,no. 4, pp. 323-326.

[11] Yadav AK, Roy R, Kumar R, Kumar CS & Kumar AP 2015, ‘Algorithm for de-noising of color images based on median filter’, 2015 Third International Conference on Image Information Processing (ICIIP), pp. 428-432.

[12] YousufMNNaMA 2011, ‘A New Method to Remove Noise in Magnetic Resonance and Ultrasound Images’, Journal of Scientific Research, vol. 3,no. 1, pp. 81-89.