A Novel Comprehensive Spectral-Textural Features Extraction of Laplacian Pyramid Enhanced Retina Images for Early Detection and Diagnosis of Diabetic Retinopathy

Omed H. Hama[1], Prof. Dr. Şükrü ÖZGAN[2]Asst.prof.Dr.Raghad Z. Yousif[3]

[1](PhD student, Physics Department/ Kahramanmaraş SütçüimamUniversit,Turkey)
[2](Physics Department,Kahramanmaraş Sütçüimam University, Turkey)
[3](Physics-Medical Physics Department, Science/ Salahaddin University-Hawler, Iraq)

Abstract: *Diabetes is the leading cause of blindness in western countries in the productive age groups. For diabetic retinopathy the major risk is the presence of diabetic. Early detection can potentially reduce the risk of blindness. The eye is one of the most important end organs for complications of diabetes. There are two types of ocular diabetic complications: direct and indirect. Diabetic retinopathy is the primary direct complication of diabetes in the eye. There are two main categories of DR, Non-proliferative (Non-PDR) and Proliferative (PDR). In this work a new technique is presented for automatic screening of diabetic Retinopathy based on Comprehensive Spectral-Textural features which are two fundamental patterns elements that can be interpreted by human. To extract spectral features parameters (Bi-spectral invariant features at different angles of orientation)the gray level retina images are subjected to Radon transform which convert the 2D image in to 1D image vector which introduced to the HOS (Higher Order Statistic) Scheme. While nineteen Textural features parameter were derived from gray level matrix, co-occurrence matrix, and run length matrix features combined to perform texture analysis for retina images and hence to generate the features vector. Then after twenty six best features form both categories have been considered and introduced to the features selection stage to reduce feature vector to 10-featurers for each retina image. The database used is in this work is MESSADOR, thus more than 150 retinal images have been used for training and testing two presented classifiers based on K-Nearest-Neighbor (KNN) and Naïve Bayes Neural Network with different rates. The simulation results show that the proposed system can achieve very high classification rates about 98.09%.*

Keywords*: CAD, Diabetic retinopathy, Radon transform, Spectral-Textural Features, KNN.*

I.   Introduction

Automatic screening of diabetic Retinopathy helps doctors to quickly identify the condition of the patient with more accurate way. Diabetic Retinopathy (DR) is asymptomatic till it becomes vision –threating. If left undiagnosed at the initial stage it can lead to partial or even complete loss of visual capacity, Retinal lesions associated with diabetes are used to evaluate different stages. The main two stages of (DR) are called non-Proliferative and Proliferative [1]. Figure (1) shows a comparison between the symptoms Non-PDR and PDR. The non-Proliferative is the earliest form where several changes of early retinopathy are seen including microaneurysms dots and some hemorrhage, cotton and wool spots, hard exudate, loops, venous bleeding and intra retinal microvascular abnormality (IRMA). These changes are mostly seen in the background and do not results in visual loss. The PDR on other hand is characterized by the proliferation of abnormal new vessels either on the disk called neovascularization on the disk (NVD) or else where the surface of retina called neovascularization elsewhere (NVE)[1][2]. Both the forms have tendency to bleed resulting in sever visual losses and thus need urgent attention for laser photocoagulation. The new vessels are combined with fibrosis resulting in the formation of fibrovasular tissue. While growing along the posterior hyaloid these leaky vessels leads to localized liquefaction of the vitreous gel. Localized collapse of vitreous gel causes vitreous contraction resulting in rupture of these fragile vessels causing vitreous hemorrhage [1]. Thus Medical imaging allows scientists and physicians to understand potential life-saving information using less invasive techniques. This automated algorithm indicates places in the image that require extra attention from the physician because they could be abnormal. These technologies are called Computer Aided Diagnosis (CAD). This paper describes components of an automatic system that can aid in the detection of diabetic retinopathy. As the number of diabetes affected people is increasing worldwide, the need for automated detection methods of diabetic retinopathy will increase as well. To automatically detect diabetic retinopathy, a computer has to interpret and analyze digital images of the retina.
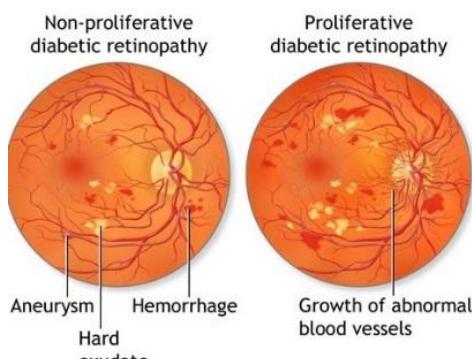

Fig1. Symptoms of Non-PDR and PDR on retina image

Extraction of, the Spectral features using Radon transform and, gray level co-occurrence matrix based Textural features from color fundus images are used. Some of related work accuracy of this work is presented in GiriBabuKande [3], whichproposes a method of polynomial contrast enhancement and dark lesion detection based on Mathematical Morphology. In this method Morphological top-hat transformation is used to segment candidate MAs from blood vessels. A.M. Mendonça et al. [4] used mean filter to the original image, obtaining a normalized image and scaling as preprocessing techniques. To discriminate microaneurysms from blood vessels "top-hat" transform and a Gaussian shaped matched filter is used. AbhirBhalerao et al. [5] used median filter for contrast normalization and contrast enhancement as preprocessing techniques. Orientation matched filter was used to differentiate microaneurysms from blood vessels. Thresholding on the output of orientation matched filter is done to obtain a set of potential candidates (MAs). Eigen image analysis applied to the potential candidate regions and a second threshold applied on the Eigen-space projection of the candidate regions eliminated certain noise artifacts. Iqbal, M.I et al. [7] used Color Space Conversion, Edge Zero Padding, Median Filtering and Adaptive Histogram Equalization as pre-processing techniques and they used segmentation to group the image into regions with same property or characteristics. Methods of image segmentation include simple thresholding, K-means Algorithm and Fuzzy C-means. AkaraSopharak et al. [8] used median filtering, contrast enhancement by Contrast Limited Adaptive Histogram Equalization and shade correction as pre-processing steps and he used Extended-minima transform for feature extraction. Priya R et al. [9] used pre-processing techniques like Gray scale Conversion, Adaptive Histogram Equalization, Matched Filter Response and proposed a method for feature extraction based on Area of on pixels, Mean and Standard Deviation. S.Saranya et al [10] use Eigenvalues from hessian matrix in detection of diabetic retinopathy. Saumitra Kumar Kuri [11] use Gabor filter with local entropy thresholding in automatic detection of (DP). Vijay M Mane etal, [12] this paper presents a unique methodology for automatic detection of red lesions in fundus images. Proposed methodology employs modified approach to matched filtering for extraction of retinal vasculature and detection of candidate lesions. Features of all candidate lesions are extracted and are used to train Support Vector Machine classifier. The next section will describe the proposed system methodology.

II. Proposed System Methodology

The framework for the proposed methodology is given in Figure 2. The input retinal images undergo preprocessing which involves image resize followed by image Denoising using median filter then image enhancements with powerful proposed Technique based on Laplacian pyramid. Then after the output from preprocessing block is introduce to both radon transform system which is start by theta selection followed by applying Radon transform before spectral features extraction using HOS (Higher order statistics technique) Scheme .The Textural features involve a features derived from Texture analysis tools, GLCM (Gray Level Co-occurrence Matrix) and run length matrix. These features were combined with spectral features to feature vector. The initial feature vector contains basically 26 features introduced to feature selection stage, which aimed to select the top 10 features from both groups. These features then will be used for training two types of classifiers (KNN and Bayes Naïve) to classify the category of retina images (normal or abnormal (PDR, or Non-PDR)) based on the presence of DR.
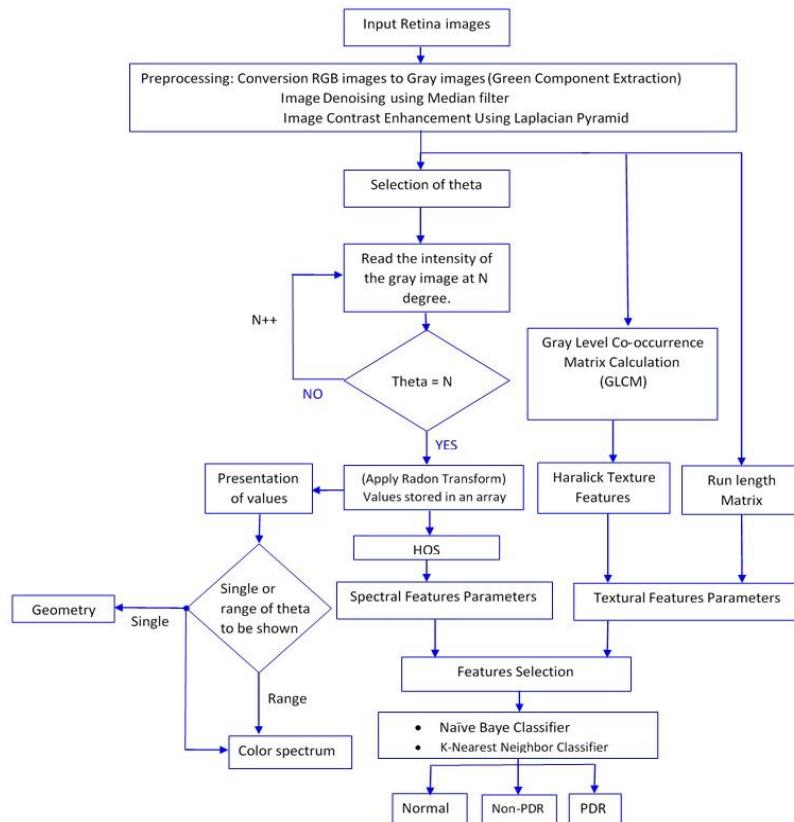

Fig 2. Proposed Methodology Flow chart

A. *Preprocessing of Image Data*

Feature extraction is most important part in this project and is widely used in classification processes. This extraction is carried out after preprocessing the images. It is thus necessary to resize retina images convers it to gray, Denise images and improve the contrast of the image, which will result in better feature extraction process. Figure 3, shows the pre-processing step primarily consists of resizing image to 512 by 512, then the green component is extracted and introduced to the Median Filter The median filter is a non-linear filter type and which is used to reduce the effect of noise without blurring the sharp edge. The operation of the median filter is – first arrange the pixel values in either the ascending or descending order and then compute the median value of the neighborhood pixels. Laplacian pyramid has been used for image enhancement so as to obtain an output image persisting its naturalness with an improved local and global contrast. Laplacian pyramid is used for separating the brightness and contrast components of an image. The brightness component is characterized by slow spatial variations and contrast components tend to vary abruptly. Therefore, the brightness component has low frequency while the contrast component tends to have a relatively high frequency. Each band of Laplacian pyramid [13][14] is the difference between two adjacent low-pass images of the Gaussian pyramid [I0, I1… IN]. Image enhancement is an important step in preprocessing stage. The objective of image enhancement is to increase the visual perception of the image o that they are more suitable for human viewers or machine vision applications. It is well known in the image processing society that there is no unifying or general theory for image enhancement algorithms. Thus an enhancement algorithm that is suitable for some application may not work in other applications. In image processing technology, image enhancement means improving image quality through a broad range of techniques such as contrast enhancement, color enhancement, dynamic range expansion, edge emphasis, and so on.

(a):original retina image   (b):resized retina image to 512*512

(c):Green Componet separation   (d):meadin filtered retina image

Preprocessing Process (firt stage)
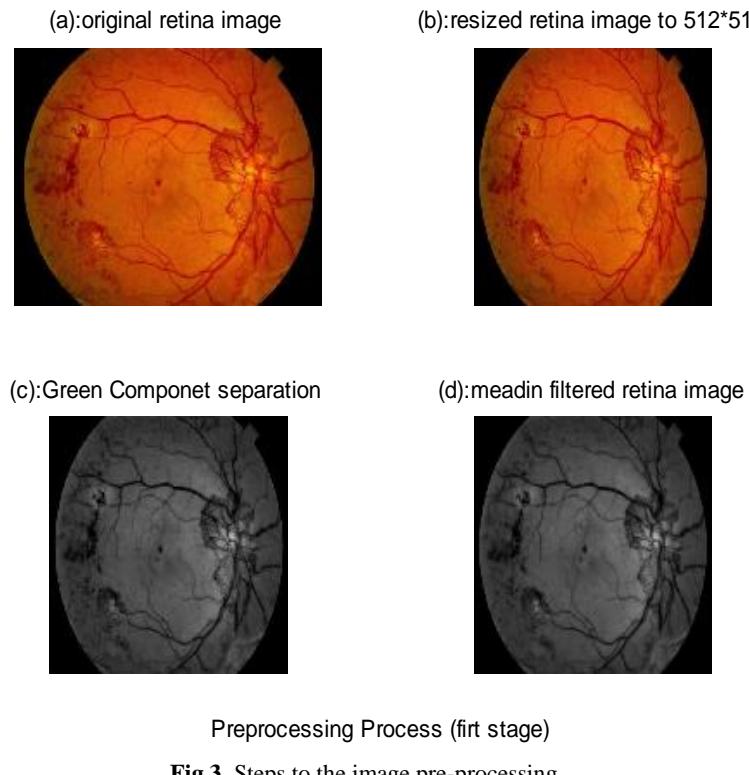
Fig 3. Steps to the image pre-processing

Thus image enhancement based on Laplacian Pyramid (LP) starts by the Generation of Laplacian Pyramid for the given image, for n particular levels. Figure 4 shows Laplacian Image decompositions for retina image. So following is the reconstruction equation

$$I_0 = I_N + \sum_{n=1}^{N} D_N \ldots\ldots\ldots(1)$$

Where N is the decomposition layer. This phase involves pyramid expansion and then its reduction. Pyramid expansion for an image of size m by n will create a result of size (2*m-1) by (2*n-1). Pyramid reduction would cause an image of size m by n to produce a result of size ceil (m/2) by ceil (n/2).

(a):I1

(b):I2

(c):I3

(d):I4
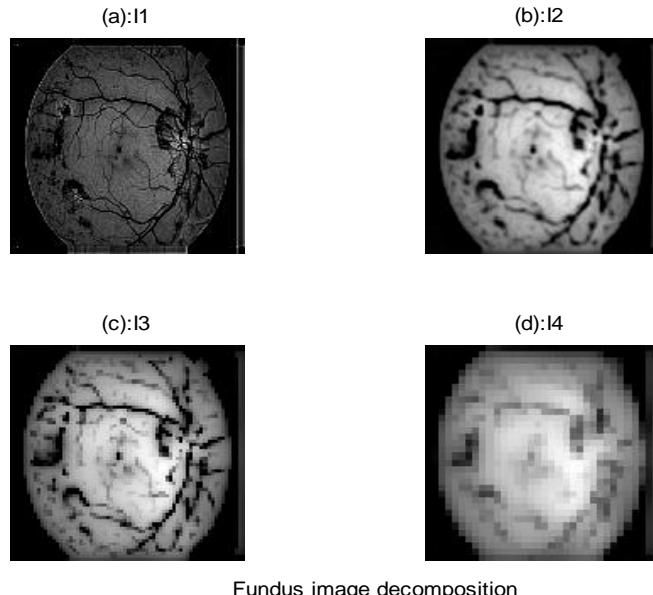
Fundus image decomposition

Fig 4. Fundus image pyramid decomposition

B. *Laplacian Contrast Enhancement*

The image enhancement technique employed by joint effort of histogram equalization and Laplacian pyramid aims to overcome the quantum jump that conventional histogram equalization based algorithms are likely to suffer from and for supplementing the improved local details that those methods cannot provide due to their inherent limitations. The proposed combination framework aims to provide an output image with well natural look without over-enhancement or severe failure and this very important to preserve the tiny areas in retina images. Contrast enhancement [15][16] improves the perceptibility o objects in the scene of enhancing the brightness difference between the objects and their backgrounds. Here the laplacian method has been used for improving contrast of an image. The histogram with gray levels in the range K [0, L-1] is a discrete function as $h(l_k) = n_k$ where l_k is the kth gray level in K and n_k represents the number of pixels having gray level l_k. In the histogram, a ridge shape with some consecutive luminance levels can be regarded as the feature area of an image. To globally distinguish between ridges and valleys and remove their ripples, we smooth the histogram [17][18][19] like as follows:

$$hg(l_k) = h(l_k) * g(l_k)............(2)$$

Where: $g(x) = -e^{x^2}$ g(x) is a Gaussian function, x is the corresponding location to a bin of the histogram, and coefficients of the Gaussian filter are normalized. Boosting minor areas, which is a key strategy of proposed contrast enhancement to suppress quantum jump. First, the peak value in the smoothed histogram $hg(l_k)$ is found as:

$$p(K) = \underset{k \in K}{Max} \{hg(l_k)\}$$ ..…………… (3)

Second, the ridges between valleys are searched and boosted. Ridge boundary is defined as the bins between the first point of the positive slope and the last point of the negative slope. The constant factor of enhancement and then the local minor areas of histogram have to be found. The local maxima has been checked, if it is found it means a peak value is found and then need to enhance it and store value to new histogram. Then Slantwise clipping stage then is started the clipping technique is used as it effectively suppresses for the quantum jump. The mean of newly generated histogram and then the mid value have been determined and then the residual from local and global clipping are gathered. Then the stage of new image generation is started, which involve finding the normalized cumulative histogram:

$$h = (cdf - cdf(min) / MN - cdf(min)) * 255.........(4)$$

Finally replace old histogram with new equalized values. Figure 5 below shows the output of proposed enhancement technique on retina image

(a):original retina image

(b):historgam of original retina image

(c):Enhanced retina image by LP

(d):Histogram of Enhanced retina image by LP
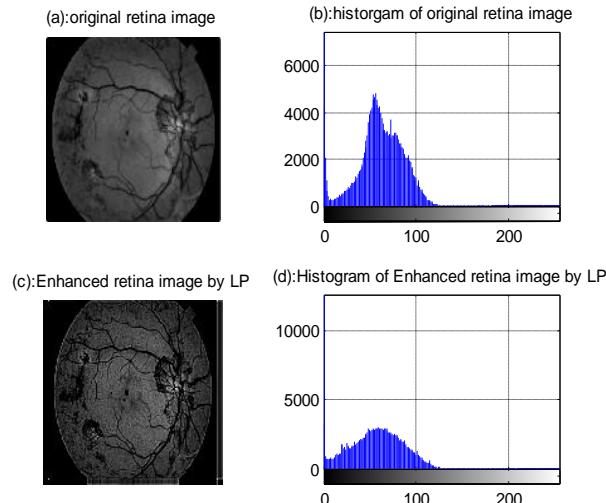
Fig5. Output of proposed enhancement technique on retina image

C. Radon Transform

Radon Transform is the projection of an image matrix along specified directions, which extract features from the image. The radon function computes the line integrals, which cut across the origin at specified angles from multiple sources along parallel paths. Each beam is 1 pixel unit apart. To convert a two-dimensional image to one-dimensional image, the radon function takes multiple, parallel-beam projections of the image from different angles by rotating the source around the center of the image. Theta is set to 0:179 where the range of projection is between zero and 179 degrees in the anti-clockwise direction. The step projection increment is default at one degree where it takes the features from the image at each and every one-degree from zero to 179 degree. The below figure explains how radon transform works. The transform function is given in equation blow:

$$R(\rho.\theta) = \int u(\rho cos\theta - Ssin\theta, \rho sin\theta + Scos\theta)ds \dots (5)$$

Where the function $R(\rho, \theta)$ is called the Radon transform of the function $u(x, y)$. This equation describes the integral along the line, s, through the image, where ρ is the distance of the line integral from the origin and theta, θ, is the angle from the horizontal [20][21][22]. The integral line extracts the features of the image where it converts the intensity of the image into values. Brighter part of the image will have a higher value while darker part of the image will have a lower value
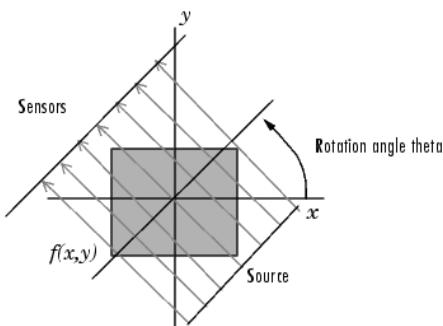

Fig 6. Projection of Radon Transform

The integral line will shift towards the origin when ρ is reduced and it can be shifted until negative p values. It will take the reading at every ρ values.
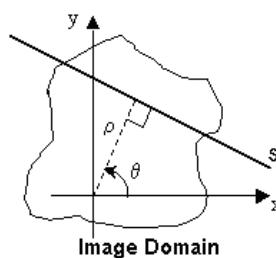

Fig7. Image domain of Radon Transform

The programmer can specify the range of the ρ values. To take the reading from another angle, the theta will change and it applies the above steps again. Moreover, it will sum up the reading values for each S value. The interception point between ρ and s is the origin of the integral line where the equation takes the reading from negative to positive infinite.

III. Feature Extraction

Feature extraction is the process of obtaining higher-level information of an image such as color, shape and texture. Spectral and textual are two fundamental pattern elements that are interpreted by human. The features vector. The feature vector used for classification consists initially of twenty-six features. These features are Homogeneity, Energy, Entropy Contrast, Symmetry, Correlation, and Momentum1, Momentum2, Momentum3, Momentum4 Angular2ndMomentum Difference, Contrast Difference, Mean Difference, Entropy Difference, Short Run Emphasis, Long Run Emphasis, Run Percentage, Gray Level Non Uniformity and , Run Length Non Uniformity .while the spectral features are ,normalized bispectral entropy1,normalized bispectral ent2, normalized bispectral entropy3 ,bispectrum phase entropy, mean bispectrum magnitude.

A. Higher Order Spectra Features

Studies have shown that the "shape" of a signal is a function of its phase signals with different wave shape and may have the same power spectrum. Assuming that there is no bi-spectral aliasing, the bi-spectrum of a real signal is uniquely defined with the triangle [23].Parameters are obtained by integrating along the straight lines passing through the origin in bi-frequency space. The region of computation and the line of integration are depicted in Figure 8. The bi-spectral invariant, p (a), is the phase of the integrated bi-spectrum along the radial line with the slope equal to a.
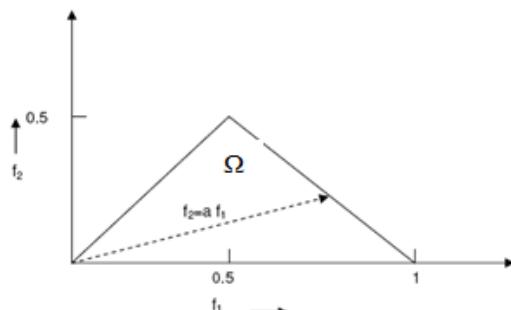

Figure 8: The Bi-spectrum curve

This is defined by:

$$p(a) = \arctan\left(\frac{I_i(a)}{I_r(a)}\right) \dots\dots\dots\dots\dots\dots(6)$$

Where

$$I(a) = \int_{f_{1=0^+}}^{\frac{1}{1+a}} B(f_1, af_1)\, df_1 = I_r(a) + jI_i(a) \dots\dots (7)$$

For 0<a≤1, and j = $\sqrt{-1}$. The variables *Ir* and *Ii* refer to the real and imaginary part of the integrated Bi-spectrum respectively. In this work, the one dimensional Radon transformed fundus images are analyzed using different higher order spectra (also known as polyspectra) that are spectral representations of higher order moments or cumulate of a signal. In particular, this paper studies features related to the third order statistics of the signal, namely the bispectrum. The Bispectrum is the Fourier transform of the third order correlation of the signal and is given by:

$$B(f_1, f_2) = E[X(f_1)X(f_2)X*(f_1 + f_2)] \quad \dots(8)$$

Where X (f) is the Fourier transform of the signal $x(nT)$ and E [.] stands for the expectation operation. In practice, the expectation operation is replaced by an estimate that is an average over an ensemble of realizations of a random signal. For deterministic signals, the relationship holds without an expectation operation with the third order correlation being time-average. For deterministic sampled signals, X (f) is the discrete-time Fourier transform (DFT) and in practice is computed at frequency samples using the fast Fourier transform (FFT algorithm). The frequency f may be normalized by the Nyquist frequency to be between 0 and 1. The bispectrum is a complex-valued function of two frequencies. The magnitude of the bispectrum may be normalized (by power spectra at component frequencies) such that it has a value between 0 and 1, and indicates the degree of phase coupling between frequency components [24]. In this study normalized bispectrum by Haubrich [25] was used to analyze the fundus images after convert them to the frequency domain by using Radon transform. Its formula is given by:

$$B_{norm}(f_1, f_2) = \frac{E[X(f_1)X(f_2)X*(f_1 + f_2)]}{\sqrt{p(f_1)p(f_2)p(f_1 + f_2)}} \quad \dots(9)$$

Where $p(f)$ is the power spectrum Bicoherence $B_{co}(f_1, f_2)$, is defined as the squared magnitude of the above normalized bispectrum. In order to differentiate the characteristics of fundus images a set of features are based on the work in [26]

have been derived. These features are the mean magnitude and the phase entropy. The features are calculated within the region defined in figure 8.
The formulae of these features are:

$$M_{ave} = \frac{1}{L} \sum_{\Omega} B_{norm}(f_1, f_2)$$
.........................(10)

$$P_e = \sum_{n} p(\psi_n) \log p(\psi_n)$$
.........................(11)

$$\psi_n = \{\phi \,|\, -\pi + 2\pi n / N \le \phi < -\pi + 2\pi(n+1) / N, n = 0,1...N-1\}$$

$$p(\psi_n) = \frac{1}{L} \sum_{\Omega} I(\phi(B_{norm}(f_1, f_2)) \in \psi_n)............(12)$$

L is the no of point within the region Ω in figure 8, ϕ Refers to the phase angle of the bispectrum, I(.) = indicator function. The bi-spectral phase is quantized and a histogram is computed as an estimate of the probability density function of the phase. The entropy used is the Shannon entropy. The mean magnitude of the bispectrum can be useful in discriminating between processes with similar power spectra but different third order statistics. However, it is sensitive to amplitude changes. Normalization can easily take care of such variation. The phase entropy would be zero if the process were harmonic and perfectly periodic and predictable. As the process becomes more random, the entropy increases. Unlike, Fourier phase, the bi-spectral phase does not change with a time shift In this work, in our attempt to characterize the regularity or irregularity of the Radon transformed retina images which contain or not the symptoms of diabetic retinopathy frombispectrum plots, Two additional bi-spectral entropies similar to that of Spectral entropy has been defined [27]. The formulae for these bi-spectral entropies are given as:
Normalized Bi-spectral Entropy (BE1):

$$p_1 = -\sum_{n} p_n \log p_n$$
...............................(13)

$$where: p_n = \frac{|B_{norm}(f_1, f_2)|}{\sum_{\Omega} |B_{norm}(f_1, f_2)|}$$
............(14)

Normalized Bi-spectral Squared Entropy (BE2):

$$p_2 = -\sum_{i} p_i \log p_i$$
...............................(15)

$$where: p_i = \frac{|B_{norm}(f_1, f_2)|^2}{\sum_{\Omega} |B_{norm}(f_1, f_2)|^2}$$
.............(16)

Normalized Bispectrum Cubed Entropy (BE3)

$$p_3 = -\sum_{j} p_j \log p_j$$
...........................(17)

$$where: p_j = \frac{|B_{norm}(f_1, f_2)|^3}{\sum_{\Omega} |B_{norm}(f_1, f_2)|^3}$$
...........(18)

The normalization in the equations above ensures that entropy is calculated for a parameter that lies between 0 and 1 (as required of a probability) and hence the entropies p_1, p_2, and, p_3 computed are also between 0 and 1.

B. Textural Features
The presented feature were derived from Gray Level Co-occurrence Matrix and run length matrix. Hence Haralick introduced Gray Level Co-occurrence Matrix and texture features [28]. It consists of two steps for feature extraction. First is to compute the co-occurrence matrix and second step is to calculate texture features based on the co −occurrence matrix. This technique has been widely used in image analysis applications especially in the biomedical field. Gray Level Co-occurrence Matrix For an image of size M x N, the gray level co-occurrence matrix (GLCM) is defined as:

$$C_d(i,j) = \left| \{(p,q), (p+\Delta x, q+\Delta y) : I(p,q) = i, I(p+\Delta x, q+\Delta y) = j\} \right|..(19)$$

Where $(p,q), (p+\Delta x, q+\Delta y) \in M \times N$, $d = (\Delta x, \Delta y)$ and $|\cdot|$ denotes the cardinality of a set. Given a grey level i in an image, the probability that the gray level of a pixel at a $(\Delta x, \Delta y)$ distance away is j is:

$$p_d(i, j) = \frac{C_d(i, j)}{\sum C_d(i, j)} \quad \text{............ (20)}$$

In this work, eleven Haralick Texture Features were derived from the gray level matrix for an image. The following features were selected. 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 neighboring pixels over the whole image. Large contrast reflects large intensity difference in GLCM.

$$Contrast = \sum_i \sum_j (i - j)^2 p_d(i, j) \quad \text{..............(21)}$$

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 (21). The addition of value '1' in the denominator is to prevent the value '0' during division. As homogeneity increases, the contrast typically decreases.

$$Homogeneity = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p_d(i, j) \quad \text{...(22)}$$

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.

Energy is sometimes derived from the use of angular second moment. It is the sum of squared elements in the GLCM known as angular second moment

$$Entropy = -\sum_i \sum_j p_d(i, j) \ln p_d(i, j) \quad \text{......... (23)}$$

$$Energy = \sqrt{Angular\ Second\ Moment} \quad \text{.......} \quad (24)$$

And

$$Angular\ 2^{nd}\ Momentum = \sum_i \sum_j p_d(i, j)^2 \quad \text{..(25)}$$

Basically, it is the measurement of the denseness or order in the image. Moments 1 to 4 are defined as:

$$m_g = \sum_i \sum_j (i, j)^g p_d(i, j) \quad \text{.......................(26)}$$

Where g is the integer power exponent that defines the moment order. Moments are the statistical expectation of certain power functions of a random variable and are characterized as follows: moment 1 is the mean which is the average of pixel values in an image; moment 2 is the standard deviation; moment 3 measures the degree of asymmetry in the distribution; and moment 4 measures the relative peak or flatness of a distribution and is also known as kurtosis.

Difference Statistics are defined as the distribution of probability $p_\delta(k)$ (k = 0,.....,n-1), and k is gray level difference between the points separated by δ in an image. They are a subset of co-occurrence matrix and the distribution of probability $p_\delta(k)$ is defined as:

$$p_\delta(k) = \sum_i \sum_j C_d(i, j) \quad \text{.......................(27)}$$

Hence the Angular second momentum (ASM) is calculated as:

$$ASM = \sum_{k=0}^{n-1} p_\delta(k)^2 \quad \text{...............................(28)}$$

When the $p_\delta(k)$ values are very similar or close, ASM is small. ASM is large when certain values are high and others are low. Contrast is also known as the second moment of $p_\delta(k)$ (its moment of inertia about the origin).

$$Contrast = \sum_{k=0}^{n-1} k^2 p_\delta(k) \quad \text{.............................(29)}$$

When $p_\delta(k)$ values are concentrated near the origin, mean is small and mean is large when they are far from the origin.

$$Mean = \sum_{k=0}^{n-1} k p_\delta(k) \quad \text{............................... (30)}$$

Entropy is smallest when $p_\delta(k)$ values are unequal, and largest when $p_\delta(k)$ values are equal. Entropy is directly

proportional to unpredictability. The above-mentioned features were calculated for δ = (0, 1), (1, 1), (1, 0), and the total mean values on the four features were taken.

$$Entropy = -\sum_{k=0} p_\delta(k) \log p_\delta(k)$$ ……………. (31)

Run-length statistics capture the coarseness of a texture in specified directions. A run is defined as a string of consecutive pixels, which have the same intensity along a specific linear orientation. Fine textures tend to contain more short runs with similar intensities, while coarse textures have more long runs with significantly different intensities[29].Run length

matrix, $p_\theta(i, j)$ records the frequency that j points with a gray level i continue in the direction θ. The (i)th dimension of the matrix corresponds to the gray level and has a length equal to the maximum gray level , n, while (j)th corresponds to the run length and has length equal to the maximum run length , l. Five measures from run -length matrixes of θ = 0̊, 45̊, 90̊ ,and 135̊ were computed by as follows.

$$Short\ Run\ Emphasis = \frac{\sum_i \sum_j p_\theta(i, j) / j^2}{\sum_i \sum_j p_\theta(i, j)}$$ ……….. (32)

Short run lengths are emphasized by dividing each run length value by the square of its length. The total number of runs in the image is the denominator.

$$Long.run.Emphasis = \frac{\sum_i \sum_j j^2 p_\theta(i, j)}{\sum_i \sum_j p_\theta(i, j)}$$ ………….. (33)

In order to allow higher weight to the long runs, each run length value is multiplied by the square of its length.

$$Run\ Percentage = \frac{\sum_i \sum_j p_\theta(i, j)}{A}$$ ……………..(34)

Where A is the area of the image. This particular feature is the ratio between the total number of observed runs in image and the total number of possible runs if all runs had a length of one.

$$Gray\ Level\ Non-Uniformity = \frac{\sum_i \{\sum_j P_\theta(i, j)\}}{\sum_i \sum_j p_\theta(i, j)}$$ …(35)

This feature depends on high run length values. The gray level non-uniformity feature will have its lowest value if the runs are evenly distributed over all gray levels.

$$Run\ Length\ Non-Uniformity = \frac{\sum_j \{\sum_i P_\theta(i, j)\}}{\sum_i \sum_j p_\theta(i, j)}$$ ……….(36)

IV. Classification

Classes based on the features the samples carry. In this project, Naïve Baye [30] and K-nearest neighbor [31][32] classifier have been employed. 150 fundus images which consist of 50 Normal fundus images and 50 Non proliferative diabetic retinopathy and 50 image Proliferative diabetic retinopathy   were used. The database used in this work is publicly available database called MESSIDOR .The MESSIDOR dataset [33]. It is composed of 1200 color fundus images acquired using a variety of retina-graphs. The image resolution is comprised between 870 pixels and 1400 pixels. All retina images has been resized to 512 pixels *512pilxels. Three different ratios are used during the training of classifiers which all depicted in Table 1

Table 1: Training Set and Ratio for testing and training.

First Training	30% Training 70% Testing
Second Training	70% Training 30% Testing
Third Training	30%Training (Random) 70% Testing

The different classes for output and training are represented as numbers shown in the table 2 below:

Table 2: Class Labels

Classes	Number
Normal	1
Non-PDR	2
PDR	3

V. Experimental and Simulation Results.

Out of 26 features retrieved with the proposed technique (Spectral-Textural), 10 features were selected using p-value, which is the probability of rejecting the null hypothesis assuming that the null hypothesis is true. The p-value has been

calculated for all presented features are subjected to the ANOVA (Analysis Of Variance between groups) test to obtain the 'p-value'. ANOVA uses variances to decide whether the means are different. Then input features of retina images have been fed into the proposed classifiers (naïve Bayes and K-Nearest Neighbor). This test uses the variation (variance) within the groups and translates into variation (i.e. differences) between the groups, taking into account how many subjects there are in the groups. If the observed differences are high then it is considered to be statistical significant. For spectral feature based system which involves (normalized bi-spectral entropy1, normalized bi-spectral ent2, normalized bi-spectral entropy3, bispectrum phase entropy, mean bispectrum magnitude), 50 bi-spectrum invariants features values were extracted for each retina image by allowing the Radon Transform orientation angle change from 0o to 80o with step of 20°.Initially Out of the 50 bi-spectrum features. Basically the lowest 'p-value' 7 features were selected. The same criteria has been applied on the textural parameters features, hence 19 features using gray level co-occurrence matrix of 135° and Run length Matrix at also 135° were selected out of 152 features (19 textures parameters by 8 angles started from 0° to 360° step 45°). Finally, from the last 26 features the best 10 were selected which have p-value less than or equal to 0.001.Then the list of features vector members used for the training of classifiers are highlighted in Table (3) below which shows the 26 nominated features from both groups (spectral and textural) and show the top 10- features which will form the feature vector.

Table (3) shows the 26 nominated features from both groups (spectral and textural) and the top 10- features.

Feature name	Feature number	p-value
Homogeneity	1	0.101006526464753
Energy	2	0.00353329582594684
Entropy	3	2.89171851479916e-13
Contrast	4	4.22363960497340e-07
Symmetry	5	0.265526907199450
Correlation	6	0.644027805784056
Momentum1	7	0.000358172821806542
Momentum2	8	0.114699863660238
Momentum3	9	0.276846890699903
Momentum4	10	0.000716929887241500
Angular 2nd Moment difference	11	0.0248040265310419
Contrast diff	12	0.114699863660238
Mean difference	13	4.22363960497340e-07
Entropy difference	14	2.45994503157047e-33
Short Run Emphasis	15	0.0289709644738788
Long Run Emphasis	16	0.405667261671826
Run Percentage	17	3.53350614552279e-10
Gray Level Non-Uniformity	18	1.85621197884877e-05
Run Length Non-Uniformity	19	0.0717542960485890
normalized bi-spectral ent1-80	20	0.0519430185057852
normalized bi-spectral ent2-80	21	1.79627385248423e-10
normalized bi-spectral ent3-80	22	0.368211707556189
bispectrum phase ent-0	23	0.000178305074482406
bispectrum phase ent-80	24	0.849087801094550
mean bi-spectrum magnitude-60	25	0.0109002773896970
mean bi-spectrum magnitude-80	26	0.513156439653202

Figure (9) depicts graphical illustration for the nominated features with their p-value; the numbers indicated in this figure were taken from table 3 (feature number).The feature vector consist the following features [Entropy, Contrast, Momentum1, Momentum4, Mean difference, Entropy difference, Run Percentage, Gray Level Non-Uniformity, Normalized bi-spectral ent2-80, bispectrum phase ent-0]. The reasons for the p-value to be very small in these seven features could be that at this parameter, there are obvious blood vessels, hemorrhages or micro aneurysms to distinguish the difference between the three groups. Figures (10) and (11) depict a graphical illustration for two out of ten features employed in feature vector.
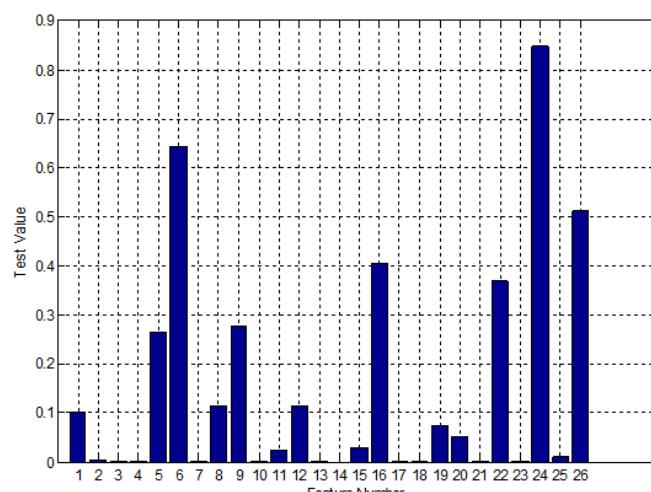

Fig 9. Depicts graphical illustration for the nominated features with their p-value
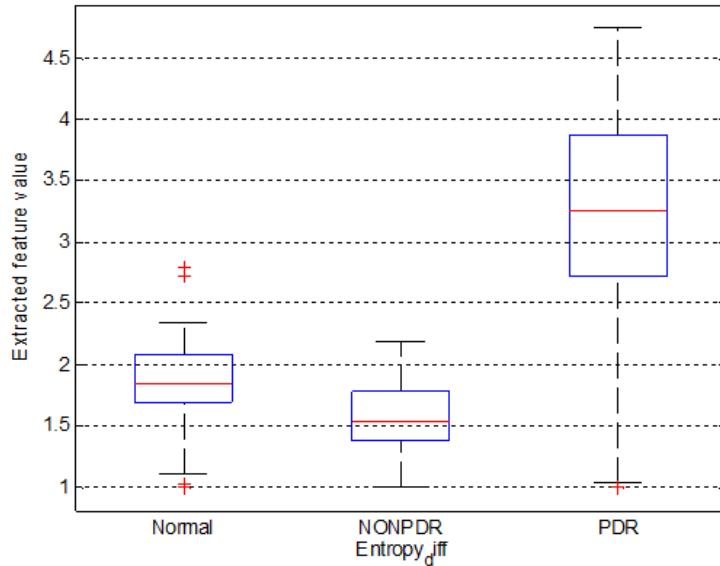
Fig10. Normalized values for the Difference Entropy feature for the three categories (Normal, Non-PDR, and PDR)
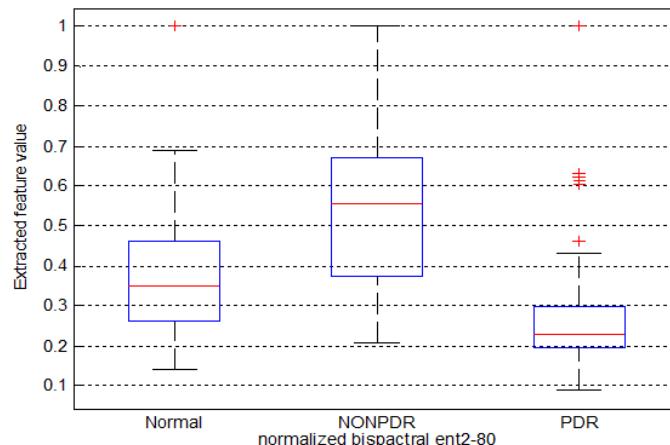

Fig11. Values for the normalized bi-spectral entropy 2 at angel theta of 80o feature for the three categories (Normal, Non-PDR, and PDR)

From Figures12, to 17 (for the training and testing files depicted in table 1) it is observed that classification using all the ten features (Spectral-Textural) features, by the Naïve Baye yields the highest classification accuracy as compared to K-nearest neighbor classifier even the fact that the proposed new technique which is a mixture between spectral and textural technique give high classification accuracy.
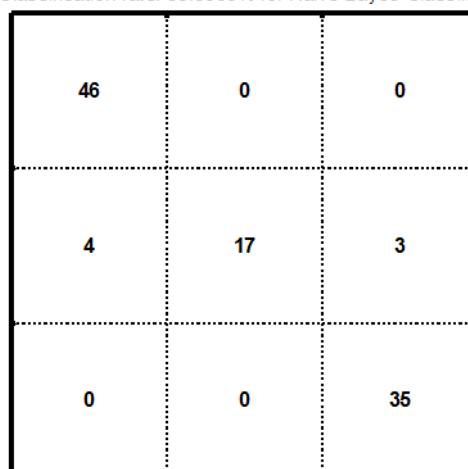
Fig 12. The resulted confusion matrix and Classification rate for first training using Naive Bayes Classifier

Classification rate: 95.5556% for Naive Bayes Classifier

11	0	0
0	17	1
0	1	15
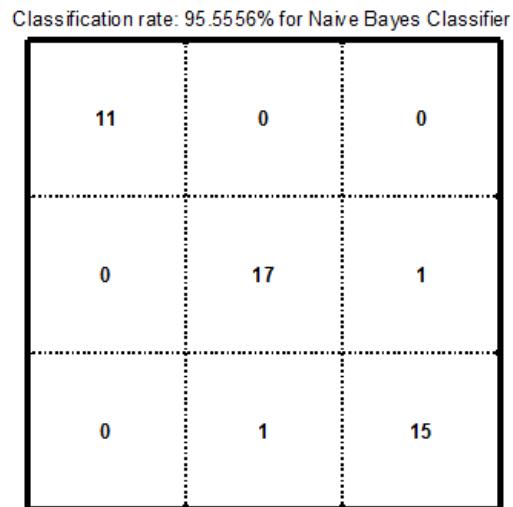
Fig13.The resulted confusion matrix and Classification rate for second training using Naive Bayes Classifier

Classification rate: 96.1905% for Naive Bayes Classifier

20	0	0
2	12	2
0	0	69
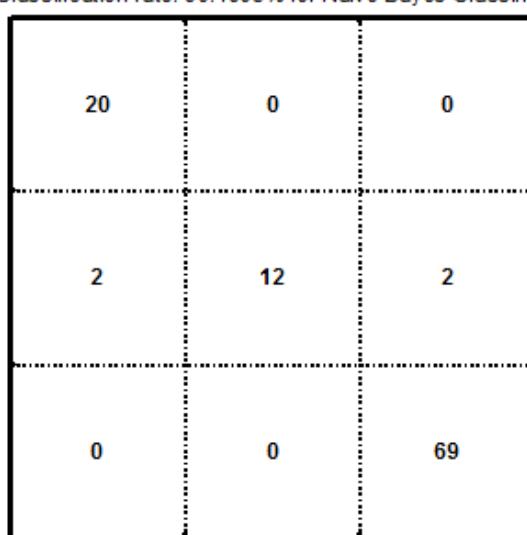
Fig14. The resulted confusion matrix and Classification rate for third training using Naive Bayes Classifier

Classification rate: 96.1905% for KNN Classifier

30	0	0
0	37	2
0	2	34
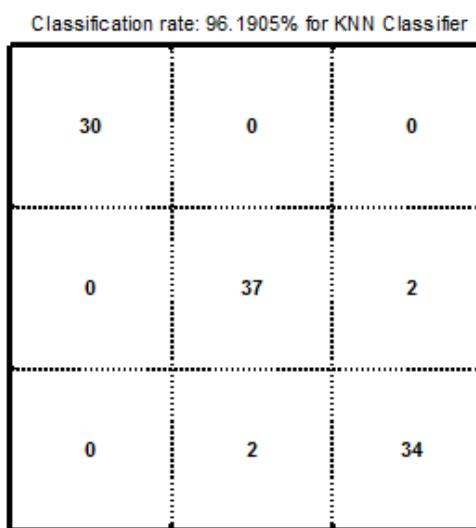
Fig15. The resulted confusion matrix and Classification rate for first training using KNN-Classifier

Classification rate: 95.5556% for KNN Classifier

11	0	0
0	19	2
0	0	13
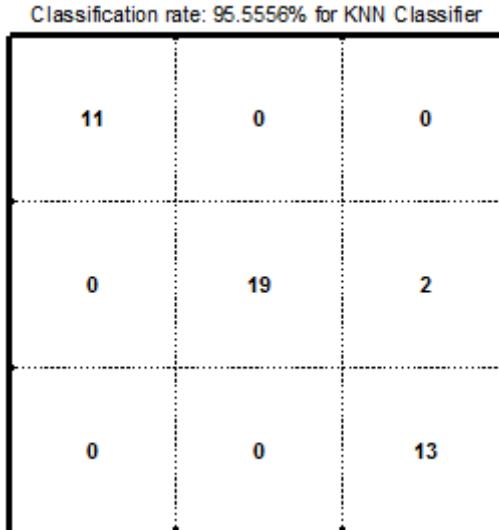
Fig16. The resulted confusion matrix and Classification rate for second training using KNN-Classifier

Classification rate: 98.0952% for KNN Classifier

41	0	0
0	11	2
0	0	51
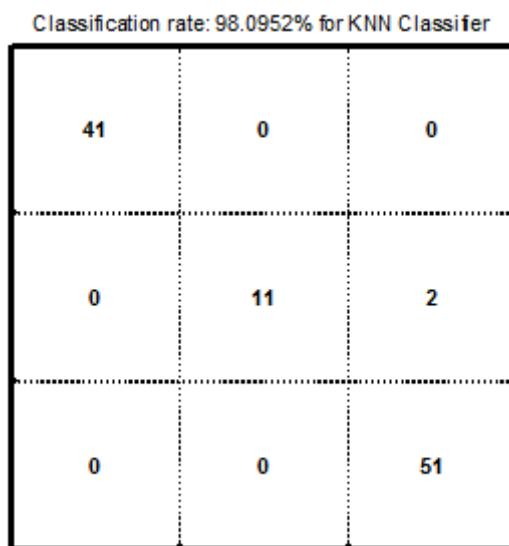
Fig17. The resulted confusion matrix and Classification rate for third training using KNN-Classifier

From Figures 12-16, it is observed that the best results for MESSIDOR database (98.095%), whereas the average was (96.6%) classification accuracy has been obtained for retina images classified with KNN classifier while using naive Bayes classifier result on an average accuracy of 97%. Thus the abnormality detection is done with higher accuracy.

VI. Conclusion

In this work, the KNN and Naive Bayes classifiers are trained through supervised learning for the features extracted to classify the retinal images. The retinal images used in this work are obtained from the publicly available MESSIDOR databases. There are two modules in this work, one that extract the image features using radon transform and one use statistical features the comprehensive features were filtered and only 10 features had been used by the proposed system The best classification accuracy obtained is 98.09%. The developed system will provide a second opinion to the ophthalmologist to do accurate diagnosis.

VII. References

[1]. VishaliGupta,etal, "Diabetic Retinopathy Atlas and Text",Jaypee Brothers Medical Publications ,First edition 2007.
[2]. Vijay M. Mane, Dattatray V. Jadhav, "Progress towards Automated Early Stage Detection of Diabetic Retinopathy: Image Analysis Systems and Potential", Journal of Medical and Biological Engineering, 34(6): 520-527.2014.
[3]. GiriBabuKande, April, 2010. Feature Extraction from Fundus Images to Analyze Diabetic Retinopathy. Research and Development Cell Jawaharlal Nehru Technological University Hyderabad Kukatpally, Hyderabad – 500 085, India.
[4]. A.M. Mendonça et al. 27 Sep 1999.Automatic segmentation of microaneurysms in retinal angiograms of diabetic patients. IEEE International Conference on Image Analysis and Processing.
[5]. AbhirBhalerao et al. 2008. Robust Detection of Microaneurysms for Sight Threatening Retinopathy Screening. IEEE Computer society.
[6]. Michael D. et al. 2010. Automated Early Detection of Diabetic Retinopathy. American Academy of Ophthalmology.
[7]. Iqbal, M.I et al. 2006. Automatic Diagnosis of Diabetic Retinopathy using Fundus images, Blekinge Institute of Technology.
[8]. AkaraSopharak et al. 2011.AutomaticMicroaneurysm Detection from None dilated Diabetic Retinopathy Retinal Images Using Mathematical using Morphology methods IAENG International journal of computer science.

[9]. Priya. R et al. 2011. Review of automated diagnosis of diabetic retinopathy using the support Vector machine, International Journal of Applied Engineering Research, Dindigul, Volume 1, No 4.

[10]. S.Saranya Rubini,etal,"diabetic retinopathy detection Based of Eigenvalues of the Hessian matrix",ElsiverProcedia Computer scinces ,47 2015 311-318.

[11]. Vijay M Mane etal," Detection of Red Lesions in Diabetic Retinopathy Affected Fundus Images", 2015 IEEE International Advance Computing Conference (IACC)

[12]. MahendranGandhi,etal," Investigation of Severity of Diabetic Retinopathy by detecting Exudates with respect to Macula" This full-text paper was peer-reviewed and accepted to be presented at the IEEE ICCSP 2015 conference.

[13]. P. J. Burt, and E. H. Adelson, "The Laplacian Pyramid as a Compact Image Code," IEEE Trans. Commun., vol. 31, no. 4, pp. 532-540, 1983.

[14]. S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitzetal., "Adaptive Histogram Equalization and Its Variations," Computer Vision Graphics and Image Processing, vol. 39, no. 3, pp. 355-368, Sep, 1987.

[15]. J. Y. Kim, L. S. Kim, and S. H. Hwang, "An advanced contrast enhancement using partially overlapped sub-block histogram equalization," IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 4, pp. 475-484, Apr, 2001.

[16]. K. Tae Keun, P. Joon Ki, and K. Bong Soon, "Contrast enhancement system using spatially adaptive histogram equalization with temporal filtering," IEEE Trans. Consume. Electron. vol. 44, no. 1, pp. 82-87, 1998.

[17]. T. Arici, S. Dikbas, and Y. Altunbasak, "A Histogram Modification Framework and Its Application for Image Contrast Enhancement," IEEE Trans. Image Process., vol. 18, no. 9, pp. 1921-1935, 2009.

[18]. S. S. Agaian, B. Silver, and K. A. Panetta, "Transform Coefficient Histogram-Based Image Enhancement Algorithms Using Contrast Entropy," IEEE Trans. Image Process. vol. 16, no. 3, pp. 741-758, 2007.

[19]. R. Fattal, M. Agrawala, and S. Rusinkiewicz, "Multiscale shape and detail enhancement from multi-light image collections," ACM Trans. Graph., vol. 26, no. 3, 2007.

[20]. S. Chandra and I. Svalbe, "A fast number theoretic finite Radon transform," 2009 Digital Image Computing: Techniques and Applications, pp. 361–368, 2009.

[21]. V. Venkatraghavan, S. Rekha, J. Chatterjee, and A. K. Ray, "Modified Radon transform for texture analysis," no. 2, pp. 9–12.

[22]. B. Kaur and M. K. Majumder, "Novel VLSI architecture for two-dimensional Radon transform computations,"2012 1st International Conference on Recent Advances in Information Technology, pp. 570–575, March 2012.

[23]. Nikias, C.L., Petropulu, A.P., "Higher-order spectra analysis: a nonlinear signal processing framework", Englewood Cliffs, N.J.: PTR Prentice Hall, 1993.

[24]. Cesar Seijas, AntoninoCaralli, Sergio Villazana," Neuropathology Classifier Based on Higher Order Spectra", Journal of Computer and Communications, 2013, 1, 28-32.

Omed Hamah Ameen Hama was born in Sulaymani/ Iraq 1980. He Received a BSc in Physics College of education, Salahaddin University/ Erbil in 2003, MSc in Biophysics in Health institution Gaziantep University/ Turkey in 2013. He started PhD in Bioengineering and Sciences in KahramanmaraşSütçüimam University in 2013 and continuous till now. Currently he is general secretary in a university in Erbil/Iraq since 2008. His field of interest is biophysics and bioengineering. He organized many international conferences, seminars and scientific competitions. He has a number of articles, research projects, theses and reports.

Prof. Dr. □ükrü ÖZGAN was born in Kahramanmaraş/Turkey 1961. He received a BSc in physics Atatürk university (Erzurum) department of education in 1984, MSc in Physics in Science of Art institution from Erciyes university/Kayseri in 1988, PhD degree in physics in Science of Art institution from Erciyes university in 1992, He became Assit professor in 2010 and currently he is professor in both Physics and Bioengineering and Sciences departments in KahramanmaraşSütçüimam university. His fields of interest are Liquid Crystals, Heat and Thermodynamics and Statistical physics. He is a supervisor of many MSc and PhD theses. He has more than 61 articles, research projects, theses and reports. He has 72 cited to his **'**SCIENCE CITATION INDEX' publications.

Dr.RaghadZ.Yousif was born in Baghdad 1975 .He Received a BSc in Electronic and Communication engineering from college of Engineering Baghdad university Department of Electronic and Communication Eng. in 1998,MSc in Electronic and Communication Engineering form department of Electrical engineering Al-Mustansriyha university –Baghdad in 2001 ,then he received a PhD degree in communication Engineering from Baghdad university of Technology Department of Electronics and Electrical engineering in 2005 .He is currently a Assistance Professor in Department of Applied physics –medical imaging in Salahaddin University also he is lecturer at Department of Computer Techniques in al Kitab private university. His Fields of interest are Network Coding Medical Imaging Swarms, Wireless Communication system optical communication system and data security. He is a supervisor of many MSc thesis and PhD thesis he was also a member of examination committee of many MSc and PhD theses .He was Awarded the prize of third best Arabic research in data security in the age of Information technology in 2003 awarded by the Association if Arab Universities.