

Machine Learning Identification of Diabetic Retinopathy from Fundus Images

Nikita Gurudath, Mehmet Celenk, and H.Bryan Riley
School of Electrical Engineering and Computer Science
Stocker Center, Ohio University
Athens, OH 45701USA

Abstract— Diabetic retinopathy may potentially lead to blindness without early detection and treatment. In this research, an approach to automate the identification of the presence of diabetic retinopathy from color fundus images of the retina has been proposed. Classification of an input fundus image into one of the three classes, healthy/normal, Non-Proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR) has been achieved. Blood vessel segmentation from the input image is achieved by Gaussian filtering. An adaptive, input-driven approach is considered for the mask generation and thresholding is accomplished using local entropy. The processed image obtained is characterized by second order textural feature, contrast, in four different orientations- 0° , 45° , 90° and 135° and structural features namely, fractal dimension and lacunarity. The research incorporates a three layered artificial neural network (ANN) and support vector machines (SVM) to classify the retinal images.

The efficiency of the proposed approach has been evaluated on a set of 106 images from the DRIVE and DIARETb1 databases. The experimental results indicate that this method can produce a 97.2% and 98.1% classification accuracy using ANN and SVM respectively invariant of rotation, translation and scaling in input retinal images as opposed to a fixed mask based on the matched filter method.

Keywords — Diabetic retinopathy, fundus images, Gaussian filtering, texture, contrast, fractal dimension, lacunarity, machine learning, artificial neural network, support vector machines.

I. INTRODUCTION

Diabetic retinopathy is a consequence to people affected by diabetes mellitus when glucose level is not kept in control [1]. It occurs as a result of an imbalance in the body's insulin level. The initial

signs of the disease are expressed in the retinal vasculature as well as in the vitreous humor (gel surrounding the retinal blood vessels). The signs occur the form of hemorrhages, exudates, cotton wool spots (CWS) and microaneurysms (MA). The presence of these abnormalities leads to NPDR. The disease progresses into a severe stage known as PDR characterized by the abnormal growth of blood vessels (neovascularization) [2-4]. The distinction between a normal retina and a retina infected with diabetic retinopathy is shown in [3].

Color fundus images captured by a fundus camera provide the input for screening of diabetic retinopathy. Fig.1. illustrates a ray diagram of the image captured by a fundus camera

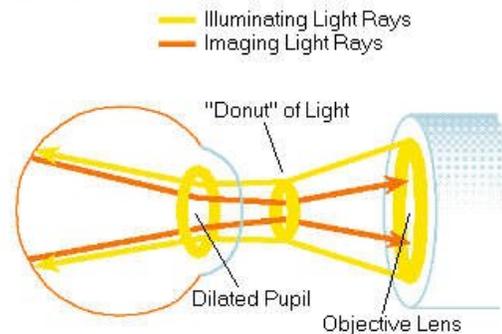


Fig. 1. Ray diagram of a monoscopic fundus image (from [5]).

At present, medical evaluation of retinopathy involves a detailed analysis of the color fundus images obtained by an ophthalmologist. The protocol followed is exhaustive and requires the support of four tests, namely, visual acuity, measurement of intra ocular pressure (IOP), gonioscopy and slit-lamp biomicroscopy [6]. It has been indicated that aforementioned tests are required since there is lack of evidence for a strong or substantial strength of support. The aim of this research is to automate the procedure to classify the input fundus image into one of the three classes by using image processing and machine learning techniques.

Review of the literature indicates extensive research is underway pertaining to the classification of diabetic retinopathy by employing image processing techniques such as thresholding, mathematical morphology and filtering [7-9]. Verma, et al., classified different stages of diabetic retinopathy utilizing six features: area and perimeter of the red, green and blue layers of the original retinal images obtaining 91% accuracy [10]. The use of fractal characteristics to classify diabetic retinopathy provides an alternative approach to deal with non-Euclidian geometry of the retinal vasculature [11]. Agurto, et al., employed the use of textural features for retinal image analysis [12].

In this research, classification of diabetic retinopathy is performed on the original retinal images as well as the images obtained after blood vessel extraction. The features include contrast for four orientations: 0°, 45°, 90° and 135°, fractal dimension and two values of lacunarity. An artificial neural network as well as support vector machines was utilized to perform classification. The remainder of the paper is organized as follows: Section II provides a description of the approach. The experimental results are presented in Section III while section IV discusses the conclusion and future work.

II. DESCRIPTION OF THE METHOD

For this research, a set of 106 images were obtained from the DRIVE [13] and DIARETDB1 [14] database. There were 30 normal images, 53 NPDR images and 23 PDR images. The number of NPDR images are higher in order to train the system to identify a class that has similarities to the other two classes. The primary research approach involves three major steps and are described in the following sections.

A. Blood vessel segmentation by Gaussian filtering and adaptive mask generation

Matched filter techniques are used to approximate the gray-level profile of a blood vessel by a Gaussian distribution [15 - 17]. The intensity profile generated from a grayscale retinal image in this research is illustrated in Fig. 2.

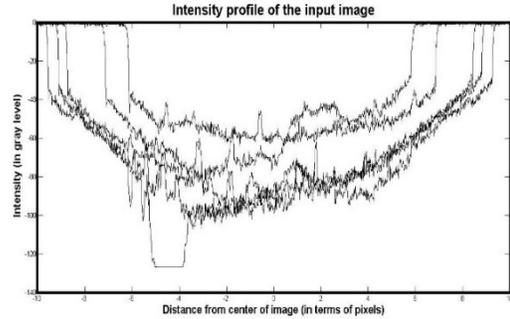


Fig.2. Intensity profile of an input image.

The intensity profile is instrumental in modelling the kernel function for filtering. Blood vessels have a tapering structure towards the edges. It is assumed that this change is gradual and a uniform width is considered [18]. Based on the information from Fig. 2, a Gaussian kernel is chosen as a smoothing function. The kernel aids in segregating the vessel edge from its background. For a grayscale input image $I(n_1, n_2)$, the Gaussian function is given as

$$G(n_1, n_2) = -\frac{1}{2\pi\sigma^2} e^{-\frac{n_1^2}{2\sigma^2}}, |n_2| \leq \frac{L}{2} \quad (1)$$

where (n_1, n_2) is the image plane axes, σ is the spread of the intensity profile and L is the length of a blood vessel segment that is assumed to be along the n_2 axis. The negative sign reflects the fact that blood vessels in a fundus image have lower reflectance compared to the other retinal surfaces. Hence, they appear darker than the background.

Since blood vessels are oriented arbitrarily, the function must be rotated in all possible directions. The angular resolution θ determines the number (N) of kernels required, which is given by

$$N = \frac{360}{\theta} \quad (2)$$

In this work, N number of kernels are convolved with the original image $I(n_1, n_2)$ of size $N_1 \times N_2$, and at each pixel (n_1, n_2) only the maximum response is elicited. The resulting image $I_g(n_1, n_2)$ is subjected to a local thresholding scheme based on entropy.

A mask $M(n_1, n_2)$ is generated in order to define the region of interest to carry out thresholding. The mask is determined depending on the grayscale input image, $I(n_1, n_2)$. A basic block diagram indicating the adaptive mask generation process is shown in Fig. 3.

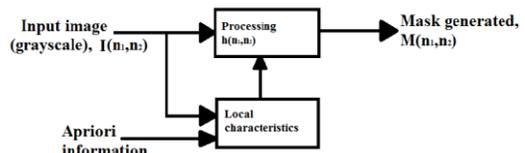


Fig. 3. General adaptive image processing system utilized to generate the mask (from [19]).

The system impulse response $h(n_1, n_2)$ is input driven and it is defined by the Gaussian kernel.

$$h(n_1, n_2) = G(n_1, n_2) = -\frac{1}{2\pi\sigma^2} e^{-\frac{n_1^2}{2\sigma^2}}, |n_2| \leq \frac{L}{2} \quad (3)$$

The mask generated is a linear convolution between the system impulse response and the input grayscale image. It is given as

$$M(n_1, n_2) = h(n_1, n_2) * I(n_1, n_2) \quad (4)$$

Local entropy thresholding entails the formation of gray level co-occurrence matrix [20] $C_{m,n,\emptyset}$ that is defined as

$$C_{m,n,\emptyset} = \sum_{n_1} \sum_{n_2} P\{I'(n_1, n_2) = m \ \& \ I'(n_1 \pm d\theta_0, n_2 \mp d\theta_1) = n\} \quad (5)$$

where d is the distance between the pixels, \emptyset is the orientation and $P\{\cdot\} = \begin{cases} 1, & \text{if argument is true} \\ 0, & \text{otherwise} \end{cases}$

The size of $C_{m,n,\emptyset}$ is specified to be $N_1 \times N_2$. Next values of θ_0 and θ_1 indicating different orientations and where $d = 3$ are provided in Table 1.

TABLE 1: Values of θ_0 and θ_1 for various \emptyset .

\emptyset	θ_0	θ_1
0°	0	3
45°	-3	3
90°	3	0
135°	3	-3

Let R denote the number of pixel pairs possible for a particular orientation, then the probability of occurrence is given as

$$P(m, n) = \frac{C_{m,n,\emptyset}}{R} \quad (6)$$

In this type of thresholding, the foreground and the background pixels are considered as different sources. If $0 \leq T \leq N-1$, then entropy for foreground pixels is

$$H_f^{(2)} = \sum_{i=0}^T \sum_{j=0}^T P(i, j) \log_2\left(\frac{1}{P(i, j)}\right) \quad (7)$$

Similarly, the entropy for background pixels is

$$H_b^{(2)} = \sum_{i=T+1}^{N-1} \sum_{j=T+1}^{N-1} P(i, j) \log_2\left(\frac{1}{P(i, j)}\right) \quad (8)$$

where the superscript, $^{(2)}$, indicates that the measure of entropy is a second order statistic. The optimum threshold T_{opt} is formulated as

$$T_{opt} = \operatorname{argmax}[H_f^{(2)} + H_b^{(2)}] \quad (9)$$

B. Feature extraction

Feature extraction is performed on the image after thresholding, $I_t(n_1, n_2)$. The nature of the fundus images are such that classification requires surface inspection. Texture of images provide information about the spatial distribution of gray levels that is integral to defining the regions in fundus images that have abnormalities [20]. The second order statistic, contrast in four orientations of 0° , 45° , 90° and 135° , have been calculated. Fractal features describe the similarity of pixels at different scales and can effectively identify the gaps in an image. Three fractal features, namely - fractal dimension and lacunarity considering foreground pixels and lacunarity considering both foreground and background pixels have been computed for all the images utilized in this research.

The co-occurrence matrix is calculated for $I_t(n_1, n_2)$ as described by eqns. (5) and (6). For an orientation \emptyset , the contrast [20] is calculated as

$$CON_\emptyset = \sum_i \sum_j P(i, j)(i - j)^2. \quad (10)$$

Objects that have integer dimensions conform to traditional Euclidean geometry. Objects possessing the property of self-similarity are known as fractals [21]. The fractal dimension of a subset of fractals, known as wild fractals, is calculated using the box count method [21]. The image $I_t(n_1, n_2)$ is located onto a grid with mesh size s . The number of grid boxes that contain a portion of the structure is described by the power law [22]

$$B(s) = \frac{1}{s^D} \quad (11)$$

where D is the fractal dimension and is given by

$$D = \frac{\log B(s)}{\log(1/s)} \quad (12)$$

Lacunarity characterizes the distribution of gaps or holes in an image and considers the textural representation of a fractal [23]. The lacunarity is calculated for foreground and background pixels.

To conclude, the set of features used for classification of the original and processed images into the aforementioned classes (i.e., normal, NPDR and PDR) are

- Contrast in 0°
- Contrast in 45°
- Contrast in 90°
- Contrast in 135°

- Fractal dimension
- Lacunarity considering foreground pixels
- Lacunarity considering foreground pixels and empty spaces

C. Classification

Classifiers such as artificial neural networks (ANN) and support vector machines (SVM) have an extensive history in relationship to biological systems. Their linear or piecewise linear approximations capabilities are particularly attractive for pattern associations or training. Additionally, these classifiers lend to the efficient and low-cost digital hardware realizations.

A three layer, feed-forward artificial neural network is selected to implement classification using the backpropagation training algorithm [24]. The number of neurons in the input layer depends on the number of features extracted. There are three neurons in the output layer representing normal, NPDR and PDR classes. The number of neurons in the hidden layer was determined empirically. The first case considered in this research is without the adaptive mask generation and with seven features as the input.

The SVM classifier has been trained methodically using one-against-one training method and the polynomial kernel for optimization using Lagrange multipliers and Karush-Kuhn-Tucker (KKT) conditions [20]. The classifiers used have been pictorially represented in Fig. 4.

The second case involves utilizing an adaptive mask generation process as described by Eqns. (3) and (4). The schematic for this case is similar to the previous case with the exception that there are only three neurons in the input layer.

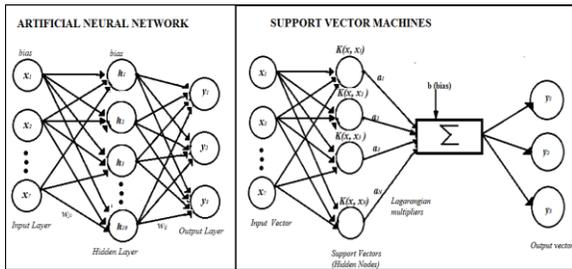
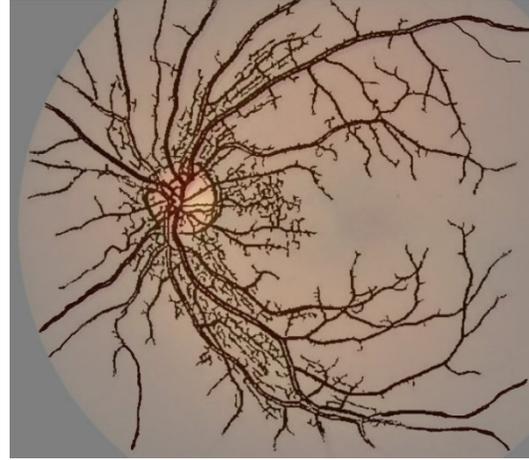


Fig. 4. Schematic diagram of ANN and SVM for case 1.

III. EXPERIMENTAL RESULTS

The original color fundus images for the three classes overlaid by the processed images are shown in Figs. 5(a), 5(b) and 5(c).



(a)



(b)



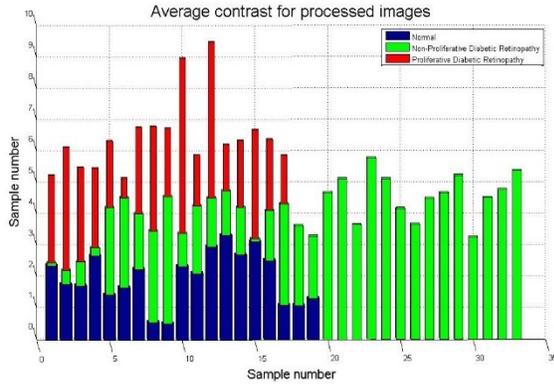
(c)

Fig.5. Processed images overlaid on top of original images. (a) Normal or healthy patient; (b) Individual diagnosed with NPDR; (c) Patient diagnosed with PDR

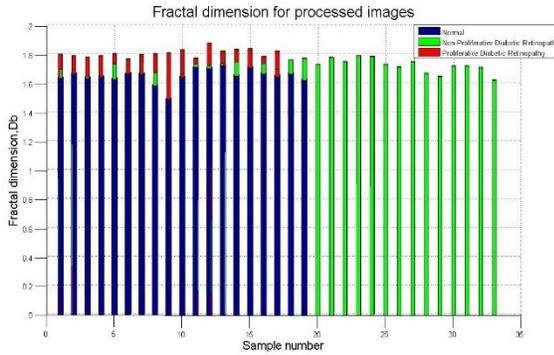
The value of L is chosen to be 9 and σ is taken to be 4. The angular resolution, θ is selected empirically as 7.5° to account for the smallest of blood vessels in

the image. The number of kernels required to obtain the processed images is 48.

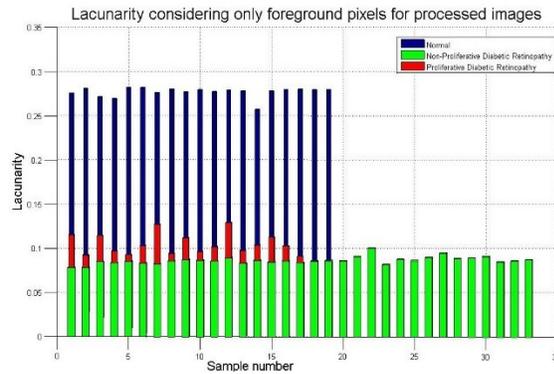
The second order statistic, contrast for $d = 3$ and for $\theta = 0^\circ, 45^\circ, 90^\circ$ and 135° , is calculated. The average contrast value and fractal features obtained for processed images is illustrated in Figs. 6(a), 6(b), 6(c) and 6(d).



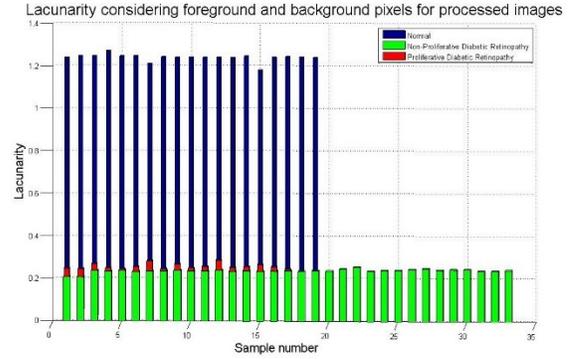
(a)



(b)



(c)



(d)

Fig. 6. Bar graph representation of features obtained for classification of diabetic retinopathy for processed images. (a) Average contrast; (b) Fractal dimension; (c) Lacunarity considering foreground pixels only; and (d) Lacunarity considering foreground and background pixels.

The feature space consisting of 7 dimensions is represented as a subset of 3 dimensions.

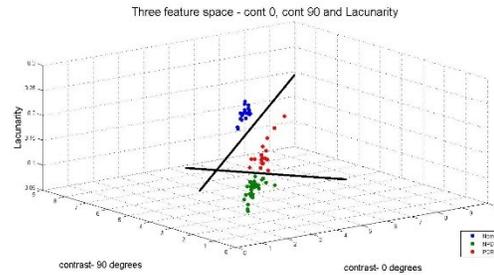


Fig. 7. Results of 3-D feature space representation with contrast for 0° and 90° orientations and lacunarity considering foreground pixels.

It is observed from Fig. 7 that a combination of the textural contrast in the horizontal and vertical direction along with lacunarity provides a distinct separation of the three classes. These three features have been utilized for classification using ANN.

Recalling the selected range of 106 samples, a two-fold cross-validation was performed varying the holdout parameter from 0.1-0.9 in steps of 0.1. In the absence of an adaptive mask generation, the classification accuracy obtained was 91.7% with ten neurons in the hidden layer. The SVM yielded an accuracy of 93% in this case. Incorporating the adaptive mask generation, the classification accuracy obtained was 97.2% using ANN with ten neurons in the hidden layer and 98.1% using the SVM.

The results of this research in comparison with other well-established results [25] are summarized in Table III. Lee et al. [26] demonstrate a classification accuracy of 82.6% for normal and NPDR and 88.3% for PDR. In [27], Nayak et al. have achieved an accuracy of 93% with sensitivity of 90% and specificity of 100%. The approach taken in [28], results in a classification accuracy of 96.15%,

sensitivity of 96.27% and specificity of 96.08%. Our research achieves an accuracy of 91.7% with a sensitivity of 93% using the ANN and an accuracy of 93% with sensitivity of 97% using the SVM when a fixed mask and seven features are used. The

classification accuracy and sensitivity obtained was 97.2% and 97% respectively using the ANN and 98.1% and 99% respectively using the SVM when the adaptive approach is utilized in conjunction with three features as shown in Fig. 7.

TABLE III. Comparative study of various DR detection algorithms

Authors	Features	Methods (Dataset size)	Salient feature	Performance measure
Lee et al. [26]	HEM, MA, exudates and CWS	NN (430)	High reproducibility	Normal-82.60% , NPDR-82.60%, PDR-88.30%
Nayak et al. [27]	Exudates, area of blood vessel and contrast	NN (140)	Texture and morphological features	Sensitivity-90% Specificity-100% Accuracy-93%
Mookiah et al. [28]	Blood vessels and exudates area, bifurcation points, global texture and entropies	GA optimized PNN classifier (156)	PNN tuning by GA and Particle Swarm Optimization (PSO)	Sensitivity-96.27%, Specificity-96.08%, Accuracy-96.15%
Our method – Case 1	Textural contrast in four orientations, fractal dimension and two values of lacunarity	NN & SVM(69)	All anomalies considered – MA, CWS, hemorrhages, exudates and neovascularization With a fixed mask	NN: Sensitivity – 93% Accuracy – 91.7% SVM: Sensitivity – 97% Accuracy – 93%
Our method – Case 2	Contrast in 0° and 90° and lacunarity considering foreground pixels only	NN & SVM(106)	All anomalies considered Adaptive mask generation procedure utilized	NN: Sensitivity –97% Accuracy – 97.2% SVM: Sensitivity – 99% Accuracy – 98.1%

IV. CONCLUSIONS AND FUTURE WORK

First order features provide less than reliable data for classification of diabetic retinopathy. Due to the nature of processed images, textural features give the necessary description that aids in robust classification. The fractal features emphasize the severity of the disease. Using a combination of the two, a classification accuracy as high as 98.1% is obtained using the SVM. This is significantly higher than that achieved by the recent methods developed in the literature. The original color fundus images are smooth in appearance. Thus, classification utilizing the features extracted from them directly does not yield high recognition accuracy. Classification using the neural net greatly depends on how well the training steps can map the data from the higher dimensional feature space to the linearly separable

classification space. As expected the processing performance depends on the number of neurons in the hidden layer.

In this research we have shown that an automatic detection of the three classes by considering all the anomalies which are critical for classification as the disease progresses. A major outcome, this research aims to check for consistency in classification accuracy when presented with a larger sample set. Considerations for future work include developing an e-health digital computer based-system that reliably implements the processing steps summarized in Section II. The commercial implementation of a certified hardware prototype could then function as an effective diagnosis tool to aid in the diagnosis of individuals in regions where access to health care is limited.

V. REFERENCES

- [1] "Global status report on non-communicable diseases 2010, Geneva," World Health Organization, 2011.
- [2] A. R. Bhansar, H.Roy Sr., "Medscape," <http://emedicine.medscape.com/article/125122>. Accessed June 26, 2013.
- [3] "Diabetic Retinopathy," <http://www.labordegroup.com/procedures/diabetic-retinopathy>. Accessed 30 June, 2013.
- [4] "Diabetic Retinopathy," <http://www.eyevision.com/diabetic-retinopathy>. Accessed 30 June, 2013.
- [5] "TopCon 3D OCT," <http://www.langeeyecare.com/eye-surgery-lasik-cosmetic/topcon-3d-oct/>. Accessed 25 July, 2014.

- [6] "Diabetic Retinopathy (Initial and Follow-up Evaluation)," International Council of Ophthalmology, http://www.icoph.org/dynamic/attachments/taskforce_documents/icodiabretinoinfu_2.pdf. Accessed 26 May, 2014.
- [7] J. J. G. Leaandro, R. M. Cesar Jr., and H. F. Jelinek, "Blood vessel segmentation in retina: Preliminary assessment of mathematical morphology and of the wavelet transform techniques," Brazilian Symposium on Computer Graphics and Image Processing, 2001, pp. 84-90.
- [8] A. Sopharak, B. Uyyanonara, S. Barman, and T.H. Williamson, "Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphological methods," Computerized Medical Imaging and Graphics, vol. 32, pp. 720-727, 2009.
- [9] A. Budai, G. Michelson, and J. Harnegge, "Multiscale blood vessel segmentation in retinal fundus images," Conference proceedings BVM 2010, Aachen, 2010, pp. 211-215.
- [10] K. Verma, et al., "Detection and classification of diabetic retinopathy using retinal images," Annual IEEE India Conf. INDICON-2011, Hyderabad, India, 16-18 Dec. 2011, pp.1-6.
- [11] S. Talue, and S. Giovanzana, "Image analysis of normal human retinal vasculature using fractal geometry," International Journal of the Bioflux Society, vol. 4, issue 1, 2012, pp. 14-18.
- [12] C. Agurto, H. Yu, V. Murray, M. S. Pattichis, S. Barriga, W. Bauman, and P. Soliz, "Detection of neovascularization in the optic disc using an AM-FM representation, granulometry and vessel segmentation," 34th Annual International Conf. IEEE EMBS, San Diego, California, USA, 28 August - 1 September, 2012.
- [13] J. J. Staal, M.D. Abranoff, M. Niemeijer, M.A. Viergener, B. van Ginneken, "Ridge based vessel segmentation in color images of the retina," IEEE Trans. Medical Imaging, vol. 23, pp. 501-509, 2004.
- [14] T. Kauppi, et al., "DIARETDB1 diabetic retinopathy database and evaluation protocol," In Proc. of the 11th Conf. on Medical Image Understanding and Analysis, Aberystwyth, Wales, 2007.
- [15] M. Al-Rawi, M. Quituishat, and H. Arrar, "An improved matched filter for blood vessel detection of digital retinal images," Computers in Biology and Medicine, vol. 37, pp. 262-267, 2007.
- [16] W. L Yun, U. R. Acharya, Y.V. Venkatesh, C. Chee, L. C. Min, and E.Y.K. Ng, "Identification of different stages of diabetic retinopathy using retinal optical images," An international journal on Information Sciences, vol. 178, 2008, pp. 106-121.
- [17] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," IEEE Trans. Medical Imaging, vol. 8, no. 3, September 1989.
- [18] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," IEEE Trans. Medical Imaging, vol. 19, no. 3, March 2000.
- [19] J. S. Lim, Two-dimensional Signal and Image Processing, Prentice Hall, 1990, pp 533-536.
- [20] S. Theodoridis, and K. Koutroumbas, Pattern Recognition, 4th ed., Burlington, MA: Academic Press 2009.
- [21] H. O. Peitgen, H. Turgens, and D. Saupe, Chaos and Fractals: New Frontiers of Science, Springer-Verlag, 1992, pp. 202-213.
- [22] H. E. Hurst, "Long-term storage capacity of reservoirs," Trans. Amer. Soc. Civil Eng., vol. 116, 1951, pp. 770-808.
- [23] B. B. Mandelbrot, The Fractal Geometry of Nature, New York, 1983, pp. 14-19, 310-318.
- [24] D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," Chemometrics and Intelligent Laboratory Systems, vol. 39, 1997, pp. 43-62.
- [25] M. R. K. Mookiah, U. R. Acharya, C. K. Chua, C.M. Lim, E.Y.K. Ng, A. Laude, "Computer-aided diagnosis of diabetic retinopathy: A review," Computers in Biology and Medicine, vol. 43, issue 12, Dec 2013, pp 2136-2155.
- [26] S.C. Lee, E. T. Lee, Y. Wang, R. Klein, R. M. Kingsley, A. Warn, "Computer classification of non-proliferative diabetic retinopathy," Arch. Ophthalmol., vol. 123, issue 6, 2005, pp. 759-764.
- [27] J. Nayak, P. Bhat, U. R. Acharya, C. M. Lim, M. Kagathi, "Automated identification of diabetic retinopathy stages using digital fundus images," J. Med. Syst., vol. 32, 2008, pp. 107-115.
- [28] M.R.K. Mookiah, U. R. Acharya, R. J. Martis, C. K. Chua, L. C. Min, E. Y. K. Ng, A.Laude, "Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: a hybrid feature extraction approach," Knowl. Based Syst., vol. 39(0), 2013, pp. 9-22.