# CLASSIFICATION OF DIABETIC AND NON DIABETIC RETINAL IMAGES USING GENERALIZED FRACTAL DIMENSION

**Bhagwati Prasad[1], Kunti Mishra[2]**

*[1,2]Department of Mathematics, Jaypee Institute of Information Technology*
*A-10, Sector-62, Noida, UP (India)*

## ABSTRACT

*A method to study the changes in the retinal vasculature of diabetic patients is proposed. Generalized fractal dimension followed by canny edge detection method is used to classify diabetic people from healthy one. Statistical analysis is also done to verify the classification results.*

***Keywords: Diabetic Retinopathy, Generalized Fractal Dimension, Probability Dimension, Retina Image, Scale Factor.***

## I. INTRODUCTION

The number of diabetic people in the world is increasing significantly day by day according to a recent report (see [1]) but most of them remain undiagnosed and they may end up with more complicated diseases like diabetic retinopathy, diabetic neuropathy etc.(refer [2] and [3]). So, it is required to develop techniques to diagnose risk of diabetes in early stages. Retinal imaging is a new advancement in technology, which helps in the early detection of eye diseases.

There are number of studies available in the literature regarding classification of retina images. These techniques include the use of linear discriminant analysis, artificial neural network, support vector machine, wavelet transform etc.(for example see [4], [5], [6] and [7]). Retinal images are non linear in nature. So, techniques from non linear dynamics and chaos theory such as correlation dimension have also been given by different authors (see [8], [9] and several references thereof). In this paper, we classify retina images of diabetic and healthy persons using generalized fractal dimension (GFD). To apply GFD, first preprocessing step is employed in which we convert given colored image into grayscale and then image segmentation is done using canny edge detection method. To implement the method, we use image processing toolbox of Matlab.

First we present some basic concepts required in the sequel.

### 1.1 Fractal Theory

Mandelbrot observed that geometry of natural objects such as trees, clouds, mountains etc. does not resemble the traditional shapes in geometry. Modeling of such objects is a difficult task in comparison to man-made objects due to their irregular, non-smooth, highly complex geometries (refer [10] and [11]). He proposed fractal geometry to deal with such type of objects. Some of the well known mathematical fractals have many features in common with the shapes found in nature. One of the interesting features is their self similarity, i.e., a repetitive patterns on smaller scales. Number of authors have studied and applied this feature for different applications (for instance, see [12-14]). The retinal images also show some kind of self similarity.

## 1.2 Fractal Dimension

The concept of fractal dimension was proposed by Felix Hausdorff after observing the repetitive pattern in fractals.

Definition 1. [10] Let *A* be a non empty subset of a Hausdorff space *H*(*X*), where (*X*, *d*) is a metric space. For each $\varepsilon > 0$, let $N(A,\varepsilon)$ denotes the smallest number of closed balls of radius $\varepsilon > 0$ needed to cover *A*. Then Hausdorff dimension of *A* is given by

$$D = \lim_{\varepsilon \to 0} \left\{ \frac{\ln(N(A,\varepsilon))}{\ln(1/\varepsilon)} \right\},$$

If the limit exists.

The fractal dimension defined above is best suited for the objects which are exactly self similar in nature. But in many practical situations, objects may not be exactly self similar. The concept of box counting dimension is used in such cases. To calculate the box counting dimension of any arbitrary set *A*, we cover the set *A* with boxes and find how the number of boxes changes with the size of the boxes. If

$$D = \frac{\ln(N(r))}{\ln(1/r)},$$

Then slope *D* gives the numerical value of box counting dimension, where, *N*(*r*) is the minimum number of boxes of size *r* required to cover the object *A*.

This concept of box counting dimension has been widely used as a diagnostic tool for different diseases like cervical cancer, brain tumors, epileptic seizures etc. (for example see ([15] and [16]). Main disadvantage with the box counting dimension is that when we calculate it for any object, any box is counted or not counted at all, according to whether some points or no points exists in the box. It does not take into consideration the number of points in the box counted. So, we still remain very far from the exact measurement of the dimension. For example, in Figure 1, we have two different shapes with same box counting dimension 1.74. Thus we still did not get a good classification.

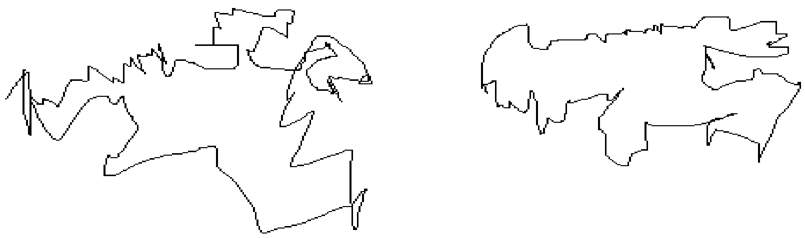


**Figure 1. Two Different Shapes with Same Box Counting Dimension 1.74**

Various other methods have also been proposed in the literature by different authors to find fractal dimension such as information dimension, correlation dimension etc. Hentschel and Procaccia [17] generalized these definitions in the form of a generalized fractal dimension (GFD) which contains almost all the above as special cases.

Definition 2. [17] Let $\mu$ be the natural probability measure on the set $A$, and $B_l(x)$ be the ball of radius l centered at a point $x$ of $A$. Then GFD is defined by $D_q$, which is

$$D_q = \frac{1}{q-1}\lim_{r\to 0}\frac{1}{\log r}\log\int d\mu(x)(\mu B_l(x))^{q-1}.$$

Our basic aim is to use GFD for the purpose of good classification of retina images to diagnose diabetic retinopathy in early stage.

## II. MATHEMATICAL ANALYSIS

First we divide the image into boxes of the size $i\times i$ for different $i$, where $i$ is a natural number, Let $B_i$ denotes the $i^{th}$ box and let $P_i = \mu(B_i)/\mu(A)$ be the normalized measure of this box. To practaically calculate generalized dimension $D_q$, we find natural probability measure, which is equivalent to the probability $p_i$, of any arbitrary point to be in $i^{th}$ box $B_i$. probability $p_i$ is calculated as number of boxes containing $i$ points divided by the maximum number of points inside a box [17].

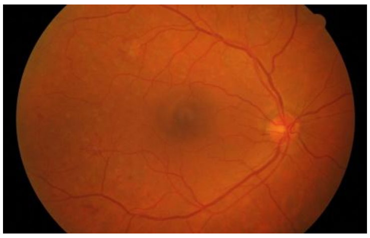
GFD for practical implementation is given by

$$D_q = \frac{1}{q-1}\lim_{r\to 0}\frac{\log\sum_{i\in r}P_i^q}{\log r}. \quad (1)$$

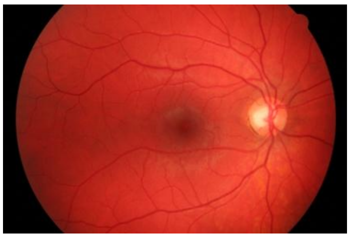
Where, $\frac{1}{q-1}\lim_{r\to 0}\log\sum_{i\in r}P_i^q$ is called generalized Renyi entropy [17]. For $q$ = 0, GFD becomes box counting dimension, for $q\to 1$, it tends to information dimension, for $q$ = 2, it becomes correlation dimension and so on. Thus, there exists infinite number of generalized dimensions of fractals. Many authors use generalized dimension for the classification in different areas like satellite imagery, texture analysis etc. (for instance see [18] and [19]). Since this method is based upon probability, it is also called probabilistic fractal dimension.

## III. DATA COLLECTION

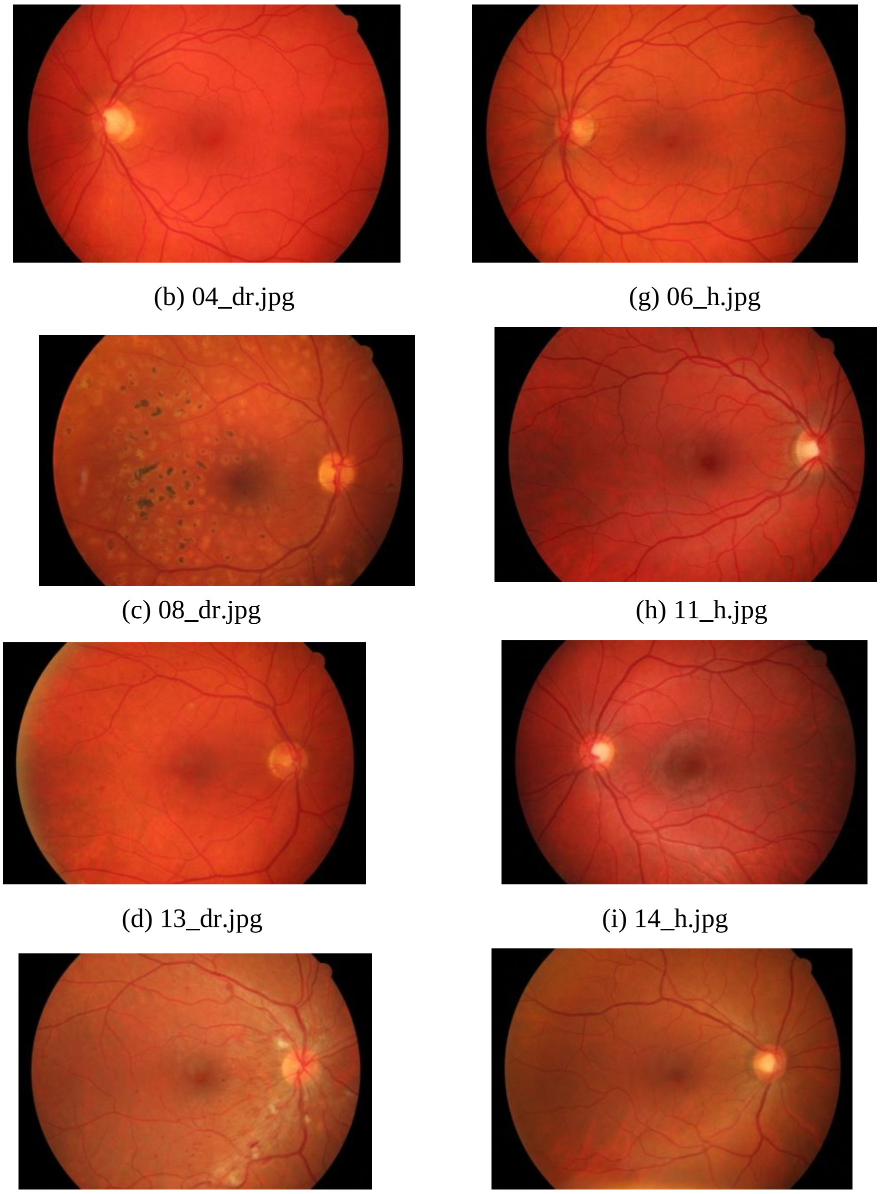
To implement the method, we use retinal image dataset available online [20].



(a) 02_dr.jpg



(f) 03_h.jpg



(b) 04_dr.jpg (g) 06_h.jpg

(c) 08_dr.jpg (h) 11_h.jpg

(d) 13_dr.jpg (i) 14_h.jpg

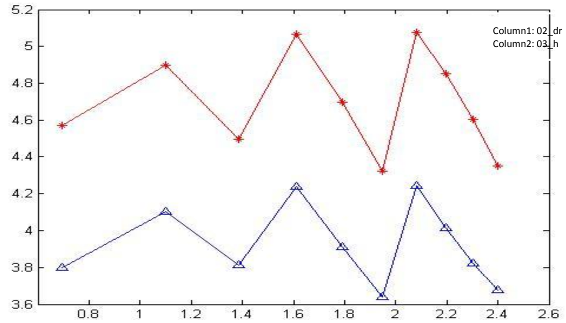
(e) 15_dr.jpg (j) 15_h.jpg

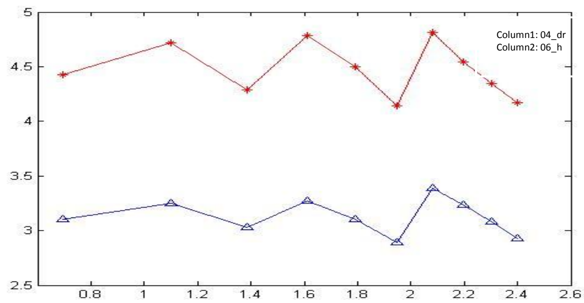
**Figure 2. (a) to (e) Diabetic Retinopathy Images, (f) to (j) non Diabetic Retinopathy Images**

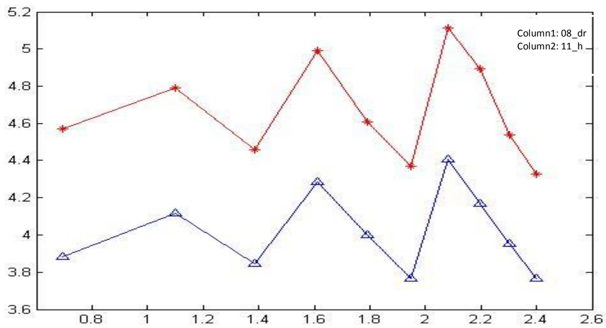
## IV. RESULTS AND DISCUSSION

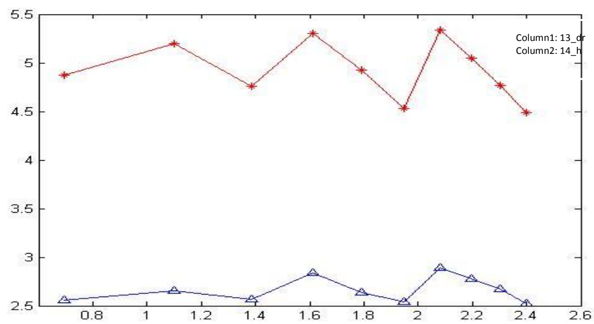
The proposed method is implemented on the selected images from the data available online ([20]). All the calculations are done by taking the value $q$ =1.8 in equation (1). To show the results, we randomly select 5 images from diabetic retinopathy images and 5 retina images of healthy eyes shown in Figure 2. The plots of generalized Renyi entropy versus log $r$ are given in Figure 3.

It is well known that the, performance of the non-parametric Kruskal-Wallis test is better than the parametric equivalent Anova test in case of asymmetric population [21]. So, for statistical verification of the present classification, we apply Kruskal-Wallis test. Further, we put GFD of diabetic data in column 1 and healthy data in column 2 for 15 images from each category and the Kruskal Wallis test is used for the two columns. Test of classification results on selected images for GFD is given in Table 1. The obtained p-values in Table 1 are very small. So, Kruskal-Wallis test also indicates that GFD gives a good classification results. We also draw box plots for the same as shown in Figure 4, which also indicate that the method serves as a good classifier for diabetic retinopathy.



ISBN: 978-81-931039-9-9

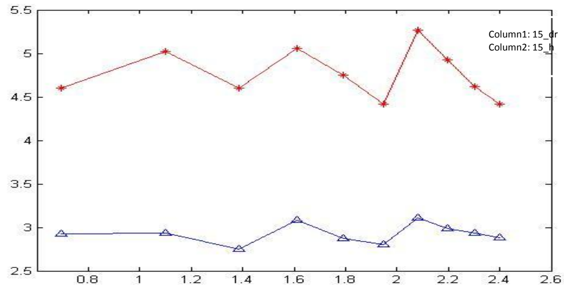
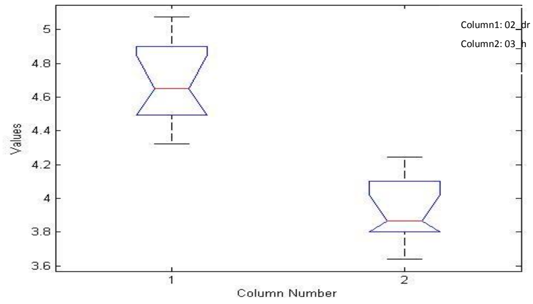
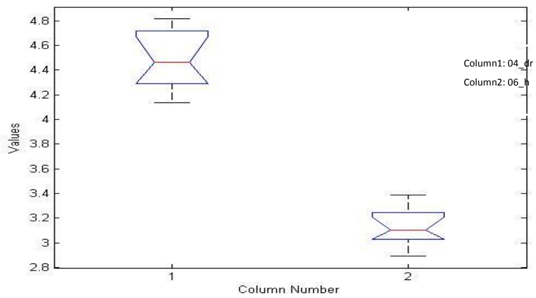
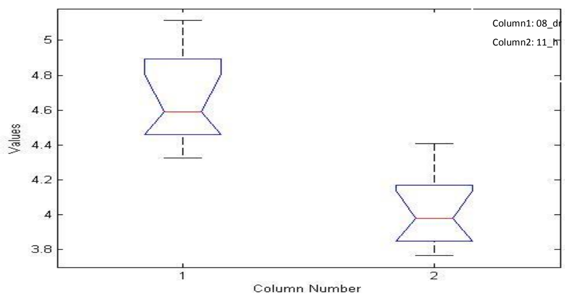
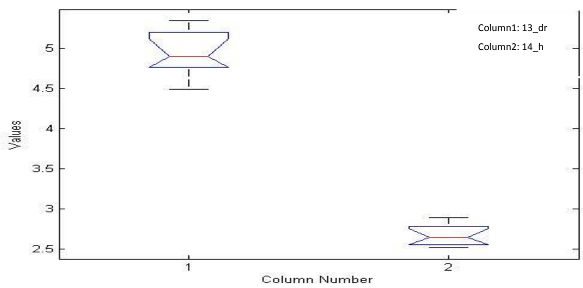



**Figure 3. Generalized Renyi entropy verus log r for both diabetic and healthy ratina images**



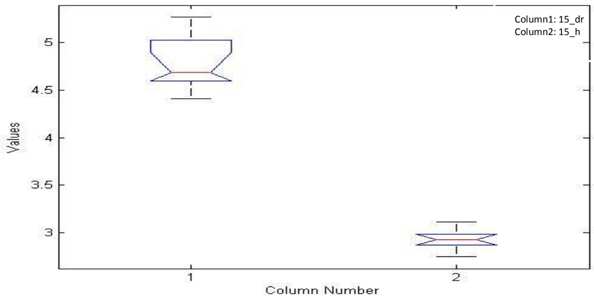



**Figure 4. Box plots for diabetic and healthy retina images**

**Table 1. Kruskal Wallis Anova Table for GFD of diabetic and healthy retina images**

| Source | SS | df | MS | Chi-sq | Prob>Chi-sq |
|---|---|---|---|---|---|
| 1. 02_dr and 03_h | | | | | |
| Columns | 500 | 1 | 500 | 14.29 | 0.0002 |
| Error | 165 | 18 | 9.1667 | | |
| Total | 665 | 19 | | | |
| 2. 04_dr and 06_h | | | | | |
| Columns | 500 | 1 | 500 | 14.29 | 0.0002 |
| Error | 165 | 18 | 9.1667 | | |
| Total | 665 | 19 | | | |
| 3. 08_dr and 11_h | | | | | |
| Columns | 460.8 | 1 | 460. | 13.17 | 0.0003 |
| Error | 204.2 | 18 | 11.344 | | |
| Total | 665 | 19 | | | |
| 4. 13_dr and 14_h | | | | | |
| Columns | 500 | 1 | 500 | 14.29 | 0.0002 |
| Error | 165 | 18 | 9.1667 | | |
| Total | 665 | 19 | | | |
| 5. 15_dr and 15_h | | | | | |
| Columns | 500 | 1 | 500 | 14.29 | 0.0002 |
| Error | 165 | 18 | 9.1667 | | |
| Total | 665 | 19 | | | |

SS=sum of squares, MS=mean of squares, DF=degree of freedom, Chi-sq=Chi-square value.

## V. CONCLUSION

We use multi scale generalized fractal dimension method for extracting descriptors to characterize diabetic retinopathy. The results verify that the GFD provides good classification which is statistically verified by Kruskal-Walli's test. Box plots and Anova tables indicate the efficiency of the method employed. This method may also provide excellent results for other retinal disease also. Moreover, this method can also be used for the classification in other areas like speech recognition, humon motion analysis, handwriting analysis, quantifying the branching frequency of virtual filamentous microbes etc.

## REFERENCES

[1]. Centre for disease control and prevention: National diabetes statistics report: estimates of diabetes and its burdern in United States. CDC, Atlanta: US department of health and human services; 2014.

[2]. J. H. Kempen, B. J. O'Colmain, M. C. Leske, S. M. Haffner, R. Klein, S. E. Moss, H. R. Taylor, R. F. Hamman, The prevalence of diabetic retinopathy among adults in the United States, arch ophthalmol, vol. 122, 2004, 552–563.

[3]. M. C. Wong, J. W. Y. Chung, T. K. S. Wong, Effects of treatments for symptoms of painful diabetic neuropathy: systematic review. BMJ, 335(7610):87, 2007.

[4]. N. F. Guler, E. D. Ubeyli, I. Guler, Recurrent neural networks employing Lyapunov exponents for EEG signals classification, Expert Sys. with Appl, vol. 29, 2005, 506–514.

[5]. V. Joshi, R. B. Pachori (2014), Classification of ictal and seizure-free EEG signals using fractional linear prediction, Biomedical Signal Processing and Control , vol. 9, 2014 1– 5.

[6]. P. E. McSharry, T. He, L.A. Smith, L. Tarassenko, Linear and non-liner methods for automatic seizure detection in scalp electroencephalogram recordings, Med. and Biol. Eng. and Comput., vol. 40, 2002, 447–461.

[7]. A. Subasi, EEG signal classification using wavelet feature extraction and a mixture of expert models, Expert Sys. with Appl., vol. 32, 2007, 1084–1093.

[8]. N. Kannathal, M. L. Choo, U. R. Acharya, P. K. Sadasivan, Entropies for detection of epilepsy in EEG, Computer Meth. and Prog. in Biomed., vol. 80, 2005, 187–194.

[9]. D. E. Lerner, Monitoring changing dynamics with correlation integrals: case study of an epileptic seizure, Physica D, vol. 97, 1996, 563–576.

[10]. B. B. Mandelbrot, Fractals : form, chance, and dimension (W. H. Freeman, 1977).

[11]. B. B. Mandelbrot, The fractal geometry of nature (updated and augmented, International Business Machines, Thomas J. Watson Research Center, H. Freeman and Company, New York, 1977).

[12]. B. Prasad and K. Mishra, Fractals in G-metric spaces, Applied Mathematical Sciences, vol. 7, 2013, 5409 – 5415.

[13]. B. Prasad and K. Mishra, A combined encryption compression scheme using chaotic maps, Cybernetics and Information Technologies, vol. 13(2), 2013, 75–81.

[14]. B. Prasad and K. Mishra, An application of IFS in encryption and compression, International Conference on Information and Mathematical Sciences, 238-241, 2013.

[15]. D. J. Barry, Quantifying the branching frequency of virtual filamentous microbes using fractal analysis, Biotechnology and Bioengineering, vol. 30(30), 2012, 1–11.

[16]. V. Rajagopalan, Liu Z, Allexandre D, Zhang L, Wang X-F, Brain white matter shape changes in amyotrophic lateral sclerosis (ALS): a fractal dimension study, PLoS one, vol. 8(9), 2013, 1–11.

[17]. H. G. E. Hentschel and I. Procaccia, The infinite number of generalized dimensions of fractals and strange attractors, Physica D: Nonlinear Phenomena, vol. 8(3), 1983, 435–444.

[18]. J. B. Florindo, A.R. Backes, M. de Castro and O. M. Bruno, A comparative study on multi-scale fractal dimension descriptors, Pattern Recognition Letters, vol. 33, 2012, 798–806.

[19]. T. Parrinello and R. A. Vaughan, Multifractal analysis and feature extraction in satellite imagery, Int. J. Remote Sensing, vol. 23(9), 2002, 1799–1825.

[20]. http://www5.cs.fau.de/research/data/fundus-images/

[21]. T. V. Hecke, Power study of Anova versus Kruskal-Wallis test, Journal of Statistics and Management Systems, vol. 15(2), 2012, 241–247.