CLASSIFICATION OF DIGITAL IMAGES USING CUMULATIVE INTERSECTION AND GENETIC ALGORITHM [+]

تصنيف الصور الرقمية باستخدم طريقة المقاطع المتراكمة والخوارزمية الجينية

Maha Abdul Amir Kadhum*   **Seham ahamed** [*]

Abstract:

A novel approach has been presented to detect and classify multiple sclerosis lesions in magnetic resonance images by using the cumulative Intersection method estimating the local fractal dimension for these images and allowing to distinguish between different types of brain tissues with respect to the human perception of roughness and statistical step - similarity of the under lying structure the offset of the Cumulative Intersection method in combination with the dimension are able to detect and distinguish lesions from other tissues by design the program Visual Basic V.6 . Also design the program to enhancement the images after reading the fractal dimension by using the genetic algorithm for resolution and high accuracy.

المستخلص:

تم عرض طريقة معدلة للكشف وتصنيف الصور الطبية ـصـور الـرنين المغناطيـسية لمـرض التـصلب المضاعف للنسيج الدماغي من خلال حساب البعد الكسوري باستخدام طريقة المقاطع المتراكمة حيث تم الفـصل بـين السطوح الخشنة والتشابه الذاتي لأنواع مختلفة من النسيج الدماغي من خلال تـصميم برنـامج Visual Basic). (V.6 بلغة كمانم تصمبم برنامج لمعالجة الصور بدقة عالية وذلك باستخدام الخوارزمية الجينية

Introduction:

Fractal Geometry is not simply on a abstract development . A cost line if measured down to its least irregularity would tend towards infinite length just as does the "Snow flake " curve . Mandelbrot has suggested that mountains , clouds , aggregates , galaxy [1] . Clusters , and other natural phenomena are similarly fractals in nature and fractal geometry application in science has become a rapidly expanding field . Besides m, the beauty of fractals has made them key element in computer graphic [2]. The fractals has are two types deterministic like Vankoch curve , Mandelbrot .Random like nature phenomena . Fractal dimension (fd) is defined as a determined value (dimension) which is independent of the scale its covering . fd is a popular parameter for describing natural textures and it represents an important feature of textural images , hence , its used to characterize roughness and self similarity in a picture . This feature is used in texture segmentation and classification shape analysis and other computer vision and image processing problems [3].

Fractal Dimension:

These are various numbers , associated with fractals which can be used to compare them . they are generally referred to as "fractal Dimension"[4]. Methods of classical geometry and calculus are not suited to study fractals and thus alternative are needed . An object normally considered as line or line segment, with dimension (D=l) , A plane , half plane or disk , surface given dimension (D=2) ;surface space or half space , or ball given dimension (D=3)is called D the topological dimension ,in order to develop a second measure of an objects dimension take a line of length(L)and divided it into (N) identical pieces each of length (l=L/N),each pieces looks like the original, with scales by ratio (r=l/N) from the whole , similarly , Two -dimensional objects , such as a square area in the plane , can be divided into "N" self -similar parts each of which is scaled down by a factor (r=l/$\sqrt{N}$) . A three dimensional object like a solid cube may be divided into "N" little cubes each of which is scaled by ratio (r=l/$^3\sqrt{N}$)with self-similarity . Thus generalization of fractal dimension is straight forward . A D-dimensional self-similar object can be divided into "N" smaller copies of itself each of which is scaled down by a factor "r" where r=l/$^D\sqrt{N}$ or N=l/r^D conversely , given a self-similar objects of "N" parts scaled by a ratio "r" from the whole its fractal or similarity dimension is given by. The D= Log (N) /Log (l/r).The fractal dimension is a non-integral dimension like Vankoch curve which have fractal dimension equal (1.2618) [5].

Main feature of fractals:

1- They have fine structure that is , they contain details at arbitrarily small scales . The more we enlarge for example ,the picture of the Mandelbrot set, the more details become apparent to eye .

2- They are too irregular to be described in traditional geometrical language , both locally and globally .

3-Often , they have from of self-similarity , perhaps approximation or statistical , the fractal dimension are greater than their respective topologic dimensions.

4- Although they are in some way , quite large set infinite , their sizes are not quantified by the usual measures such as length or area or volume as in the traditional Euclidean shapes[6] .

Cumulative Intersection Analysis:

This method, like the mass-radius method, uses the circles randomly generated within the fraction of the radius of image. The cumulative intersection method was developed by Schierwagen (1990), based on the method described by Sholl (1953). For each centre, the radius of the circle is increased by one pixel at each step. The number of separate branches is counted for each radius. This is done as follows. The points are sorted by radius and by angle for points of equal radii. The sorted array is then traversed in order and if two consecutive points of the same radius are not adjacent, the intersection count for that radius is incremented see Fig.(1). Adjacent points, including points on the diagonal, are those points touching each other.
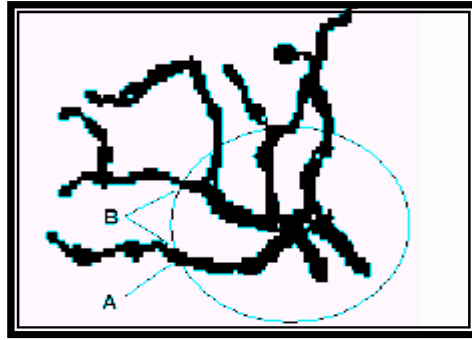
Fig.(1) A small section of image, with a single circle corresponding to the radius selected at some point during analysis

Computationally, adjacency holds when the sum of the absolute differences of the *x* and *y* coordinates respectively is less than or equal to 2. Note, that adjacent points will be part of the same branch and hence, an intersection should not be counted[6]. Consecutive points that are also adjacent are indicated by A, while consecutive points that are not adjacent are indicated by B. The analysis cases when the edge of the image is reached and no branches are found. The cumulative number of branches is calculated. The slope of the log-log relationship between cumulative number of intersections and radius is calculated by linear regression, and the fractal dimension obtained as previously described.

Convex hull mass radius analysis:

The convex hull method uses a convex hull instead of a set of circles. This involves initial computation of a convex hull bounding the extremities of the cell image Fig. (2). The hull is then scaled, so that at each step, the size of the hull is increased by one pixel, and the number of pixels incident with this hull are counted. The convex hull method is an attempt to handle self-affine fractals in a more appropriate manner. Fig. 2. Showing a convex hull (dotted line) of the image[7].
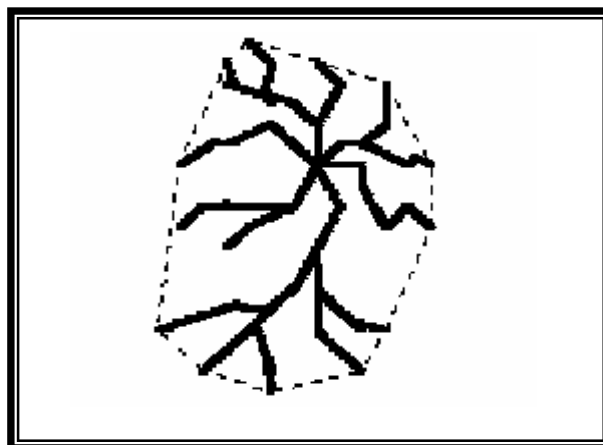

Fig. (2) Showing a convex hull (dotted line) of the image.

Vectorised Intersections:

The vectorised intersection method is an adaptation of the cumulative intersection method to vector images. The input raster image is first vectorised producing a thinned image, and a series of nodes where component lines meet. From these nodes, line segments are extracted and removed recursively starting at terminal nodes. Terminal nodes have only one line branching from them. This process is repeated until no nodes, and hence lines, remain. Analysis is then performed on a line-by-line basis, with each line segment falling into one of two analysis categories with respect to the centre point. The first category applies to lines whose end-points subtend an angle of less than 90 degrees to the centre point Fig. (3a). In this case, there exists only one point of intersection between the line and any circle with a radius between the radii of the two end points. The second category applies to lines whose endpoints subtend an angle of greater than 90 degrees to the centre point fig. (3b). The radii of the two end points are calculated, as well as the perpendicular distance of the centre point from the line. This distance will correspond to the circle of minimum radius that intersects the line. Between this minimum radius and the smaller of the two end point radii, they count two intersections. For radii between the two end point radii, they count only one intersection. The number of intersections is acquired in a similar way to the cumulative intersections method, and the stopping criteria are the same[8].
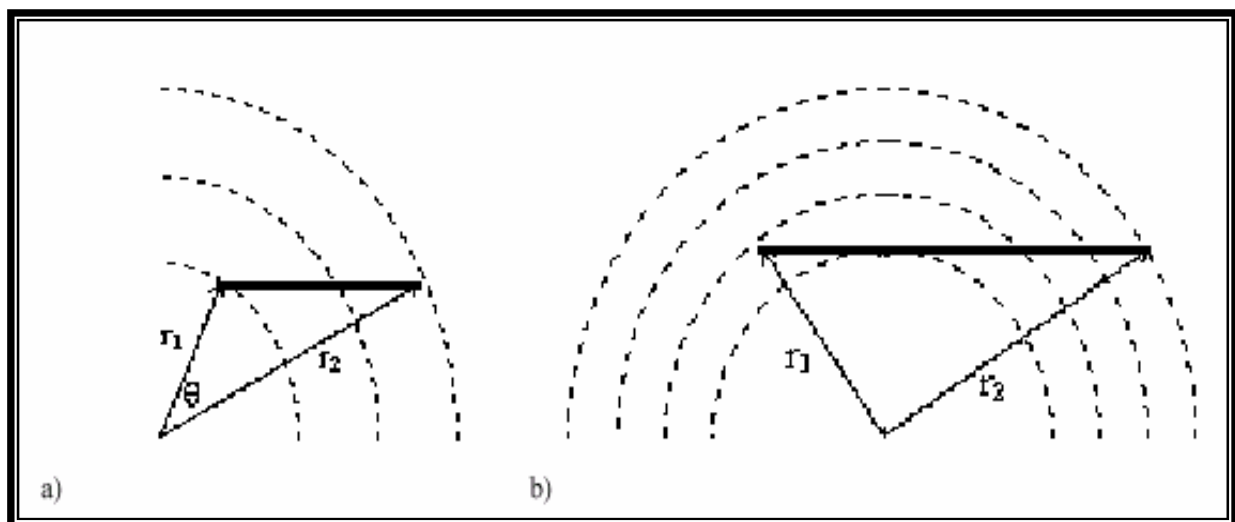
Fig.(3) The vectorised intersection method(a) Line segments whose end-points subtend an angle of less than 90 degrees count as one intersection,(b) Line segments whose end-points subtend an angle of greater than 90 degrees count as two intersections.

Genetic algorithms:

Genetic Algorithms(GA) are procedures based on the principles of natural   selection and natural genetics, that have proved to be very efficient searching for approximations to global optima in large and complex spaces in relatively short time[9]. Genetic algorithms are search procedures based on the mechanics of natural selection and natural genetics .The GAs work as follows Fig .4 :(first code each individual in the search space as a finite-length string chromosome, which consists of the characters genes (1's and 0's .A set of the chromosomes is a population .Second evaluate each chromosome with the fitness by an objective function

)fitness function .(Third apply basic operations to the population of the chromosomes. The basic operations compose of selection, crossover, and mutation .Selection is an operation which selects the chromosomes according to their fitness values .The higher the chromosome's fitness value is, the higher its probability to produce offsprings into the next generation is .Crossover creates new chromosomes by swapping genes of parent chromosomes for each pair of selected chromosomes .Mutation changes the gene of the chromosomes with a probability (mutation probability, 1 to 0 and 0 to 1) [10].

1.The basic components of GA are:

* genetic operators (mating and mutation)
* an appropriate representation of the problem that is to be solved
* a fitness function
* an initialization procedure[11].

2.Basic components, a GA works as follows.

starts by using the initialization procedure to generate the first population .The members of the population are usually strings of symbolschromosomes that represent possible solutions to the problem to be solved .Each of the members of the population for the given generation is evaluated, and, according with its fitness, it is assigned a probability to be selected for reproduction.Using this probability distribution; the genetic operators select some of the individuals .By applying the operators to them, new individuals are obtained .The mating operator selects two members of the population and combines their respective chromosomes to create offspring .The mutation operator selects a member of the population and changes some part of its chromosome .The elements of the populations with the worst fitness measure are replaced by the new individuals .[12]

```
┌─────────────────────────────────────────────┐
│ Initialize a bast population of chromosomes   │
│                                               │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│ Determine the fitness of each chromosome      │
│                                               │
└─────────────────────────────────────────────┘
                      │
                      ▼
            ╱─────────────────────╲
           ╱   Satisfy the ending   ╲          ┌──────────────────┐
          ╱        condition         ╲────────▶│ Obtain the best  │
          ╲                          ╱         │    solutions     │
           ╲                        ╱          └──────────────────┘
            ╲──────────────────────╱
                      │
                      ▼
┌─────────────────────────────────────────────┐
│ Reproduce the chromosomes According to the    │
│ fitness value                                 │
│ According to the fitness value                │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│ Recombine new chromosomes by Crossover and    │
│ mutation                                      │
│ Crossove r and mutation                       │
└─────────────────────────────────────────────┘
```
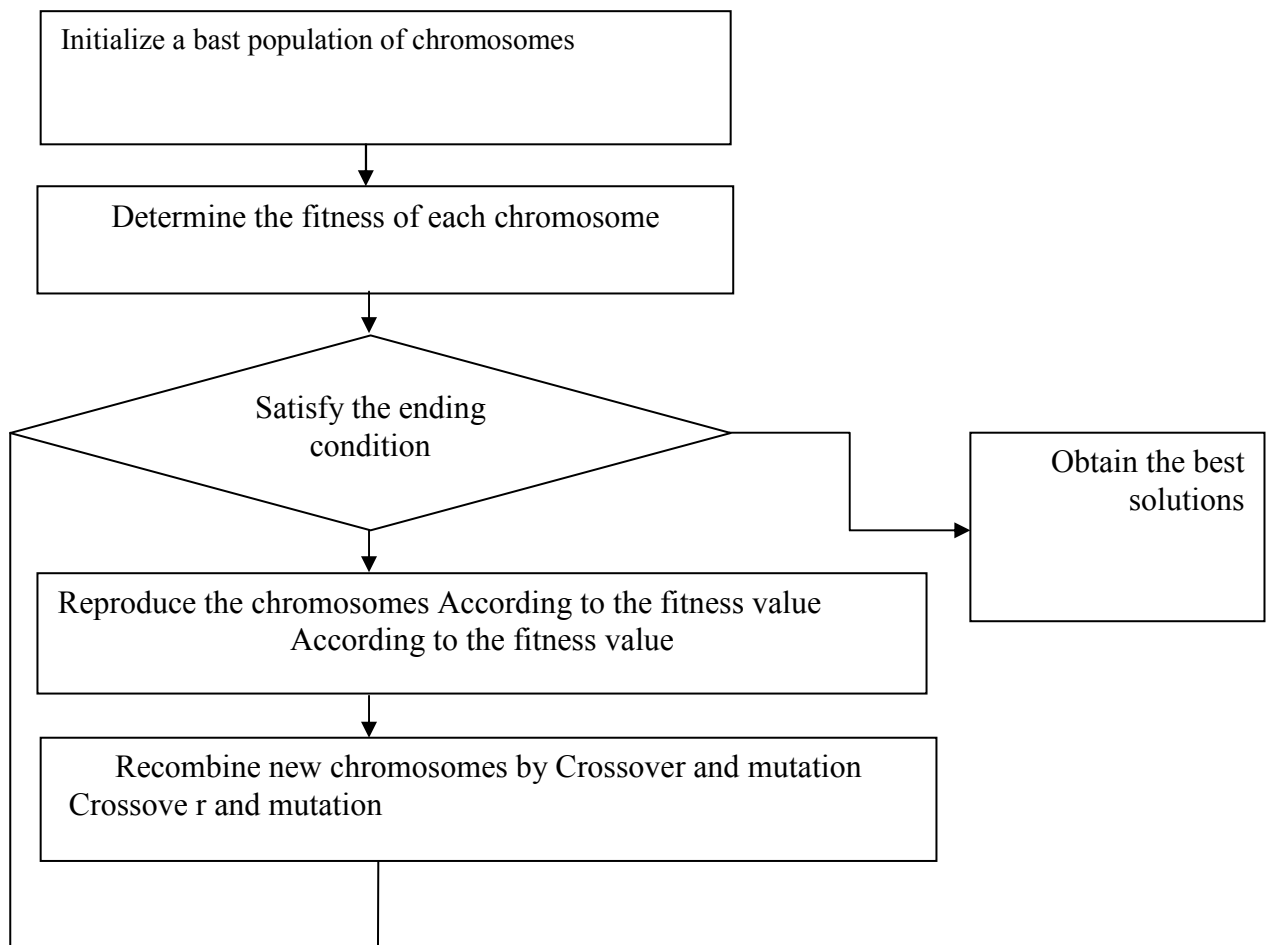
Fig(4) step finding the by solution by GA

Research Procedure:

The resulting images were classified by inspection, and divided into groups based on stratification inner (i) or outer(a)

Table(1)Showing a convex hull

	type		
strat	α	ō	Total
a	27	10	37
i	44	4	48
Total	71	14	85

part of the inner plexi form layer of the retina) and type □or d) as summarized in (dotted line) of the image. The cumulative intersection was used to calculate fractal dimension using fractal brownie motion The parameters used in method were radius factor = 0.33   and

number of centers = 20. . The Chromosome Encoding Structure Since the *CT* matrix is a signed real-valued matrix, the real value encoding is used for chromosomes representation . Since the *CT* is a 3×3 matrix, the chromosome should contain nine genes .The *CT* matrix and its chromosome encoding that is used in this work is given in Equation(1) and (2).

$$CT = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \qquad \text{.... (1)}$$

CT_Chromosome$= [C_{11}\, C_{21}\, C_{31}\, C_{12}\, C_{22}\, C_{32}\, C_{13}\, C_{23}\, C_{33}]$ ..... (2)

It must be noted that the encoding shown in equation (2)is performed via a double vector and not a binary string. For more information on how to use double vectors as GA chromosomes the reader is referred to .The Fitness Functions are applied to assess the effectiveness of each GA generated *CT* matrix which is denoted as GA .Fitness function is formulated based on the theoretical and mathematical knowledge Root Mean Squared Error. Let f be an M×N image and f' (x,y) is the corresponding reconstructed image after compressing and decompressing of image f (x,y), then the MSE is given by

$$e = \sqrt{\frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} [f(x,y) - f'(x,y)]^2}$$

........(3)

The fitness function assigns to each individual in the population a numeric value that determines its quality as a potential solution.
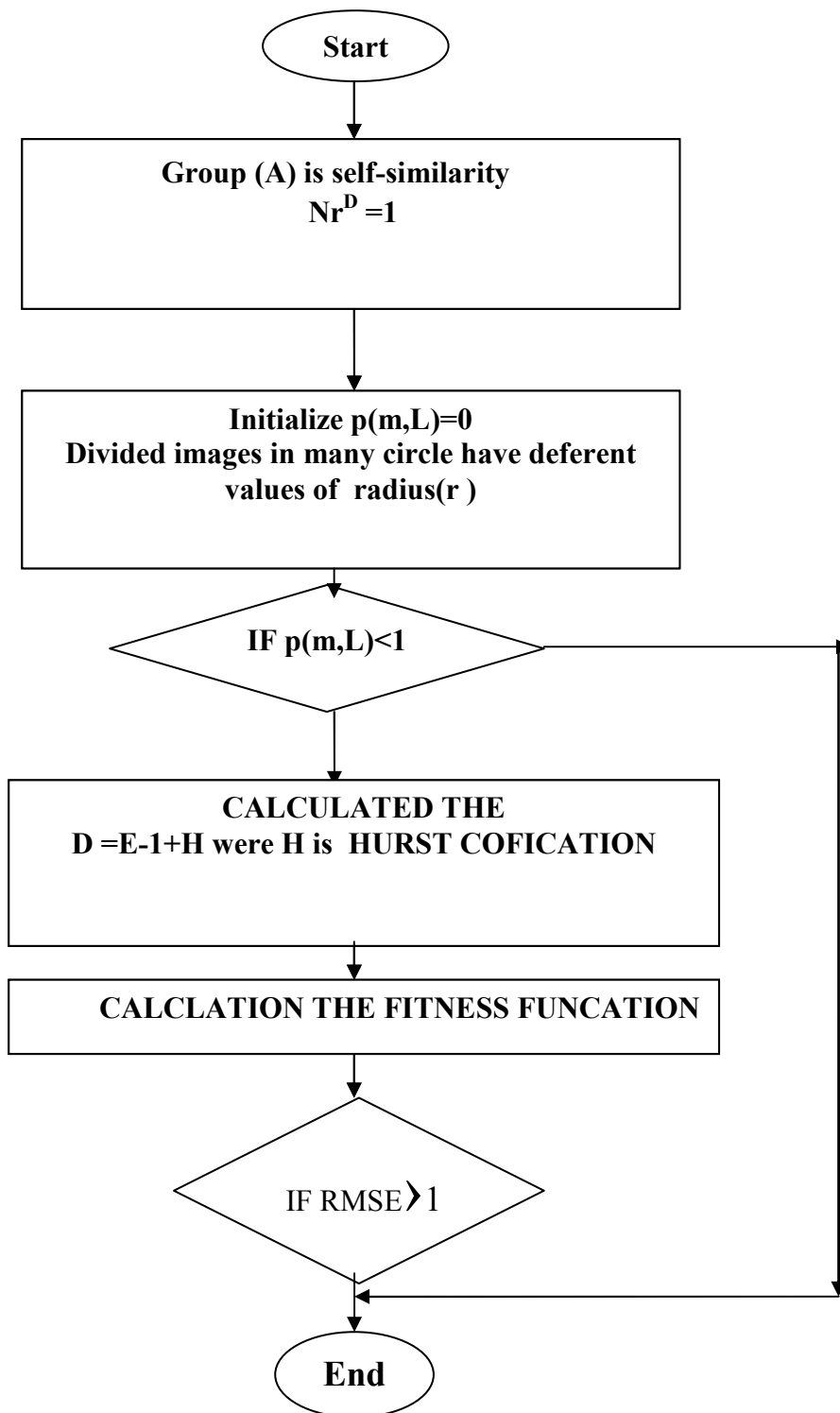
```
                        ┌─────────────┐
                        │    Start    │
                        └─────────────┘
                               │
                               ▼
              ┌────────────────────────────────────┐
              │      Group (A) is self-similarity   │
              │              $Nr^D = 1$             │
              │                                     │
              └────────────────────────────────────┘
                               │
                               ▼
              ┌────────────────────────────────────┐
              │        Initialize p(m,L)=0          │
              │  Divided images in many circle have │
              │       deferent values of radius(r)  │
              └────────────────────────────────────┘
                               │
                               ▼
                      ◇ IF p(m,L)<1 ◇ ───────────────┐
                               │                      │
                               ▼                      │
              ┌────────────────────────────────────┐  │
              │           CALCULATED THE           │  │
              │  D =E-1+H were H is HURST COFICATION│  │
              └────────────────────────────────────┘  │
                               │                      │
                               ▼                      │
              ┌────────────────────────────────────┐  │
              │  CALCLATION THE FITNESS FUNCATION  │  │
              └────────────────────────────────────┘  │
                               │                      │
                               ▼                      │
                      ◇ IF RMSE ＞1 ◇                 │
                               │                      │
                               ▼◄─────────────────────┘
                        ┌─────────────┐
                        │     End     │
                        └─────────────┘
```

Fig (5) show the research procedure

Result and discussion :

The data image (32*32) were determined of MR image for multiple sclerosis (MS)about 85patients , the gray level of these images from 0 to 255 in bitmap files a pixel pair number limit of r =30 with respect to the small neighborhood size 7*7 The program was designed and

implemented in visual Basic .6 to calculate the fractal dimension , using the cumulative Intersection method Fig (6) shows the samples of images that uses in research,
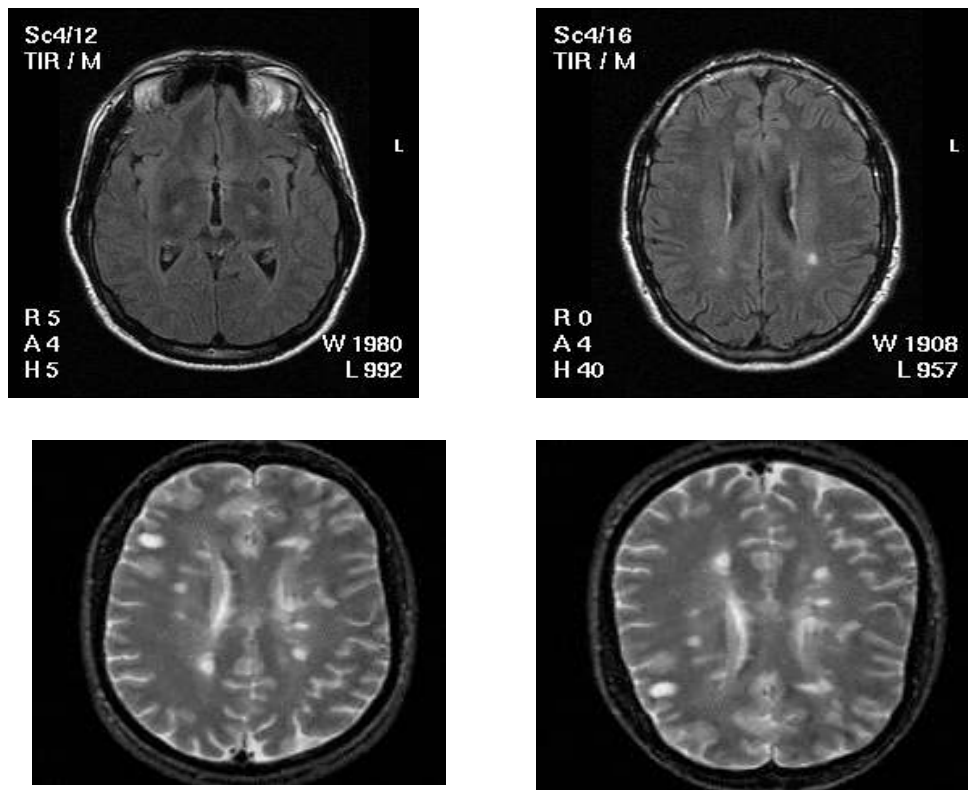

Fig (6) shows images that uses in research

And from result in table (2) The differences between all the vascular depending images fractal dimension and the "diver's sports" were highly significant ($p<0.001$) conversely the differences between the vascular bed spatial distribution and the ischemic lesions images was not statistically different .This allows us to be sure that the fractal dimension is a good tool to be used in our experimental paradigm.

The non-vascular brain lesion fractal dimension was not statistically different of the "diver's spots" one, thus our assumption was to postulate that those spots are not clearly to be defined likes vascular related ischemic lesions as generally admitted The best results are obtained for Lmax =20 .It is small, and correspond to the high frequency content of the textures .The values of fractal dimension (D)increase together with effected tissue type . The fitness denotes the individual ability to survive and to produce offspring. the fitness is the number of regions that can be coded with RMS error less than a fixed value($e >1$) .The RMS is the distance between the region and the domain block determined by its coordinates and The proposed algorithm has been evaluated on various images with different sizes .The following results are obtained for 32x32image. We present the obtained results for different configurations of error limit, number of iterations and population size.

Table (2) shown the result of image

L_{max}	Tissue Type	D unaffected			D effected		
		1	2	3	1	2	3
10	Lesion	2.31	2.35	2.32	2.61	2.60	2.63
	White	2.67	2.64	2.63	4.30	4.10	4.26
15	Lesion	2.42	2.44	2.35	2.66	2.68	2.72
	White	2.69	2.65	2.63	3.96	3.89	3.92
20	Lesion	2.31	2.32	2.32	2.61	2.60	2.62
	White	3.10	3.01	3.03	4.10	4.02	4.11
30	Lesion	2.35	2.31	2.36	2.81	2.78	2.68
	White	3.30	3.60	3.70	4.30	4.35	4.37

Conculation:

Medical images often contain branching structures, especially those obtained from neural tissue. Neurons are known to fall into several types, but distinguishing these types is a continuing problem. Automatic classification of neuron brain suitable measurements or features. Recently fractal dimension has been suggested as a useful feature. The fractal dimension is a measure of the complexity and self-similarity of an image, and is becoming accepted as a feature for automated classification of images having branching structures. Program designed to assist in the analysis of such images to calculate the fractal dimension of an image using cumulative intersection technique ,the structural attributes, and classifying these with the aid of artificial learning genetic algorithm .

Reference:

1- Mandelbrot, B. B., " Fractal Geometry of Nature", New York, W. H. freeman and Co., 1982.

2-Edgar, GA., "Measure Topology and Fractal Geometry", New York, springer-verlag, 1990.

3-Dewdency, A. K., "How to transform flights of fancy into fractal flora or fauna", Scientific American, May, 1990.

4-Penn, A. and Loew, H., "Estimating Fractal Dimension of Medical Images", http://www. Seas. Gwu. Edu/Index html, June, 2008.

5-Barnsley, M.F. and Hurd, L.P., "Fractal Image Compression", AK peters Ltd, 1993.

6-Falconer, K., "Fractal Geometry: Mathematical Foundations and Applications", John Wiley & Sons Ltd., 1990.

7-Barnsley, M. F., "Fractals Every where". Academic press, 1988.

8-Voss, R., "Random Fractals: Characterization and measurement in scaling phenomena in disordered systems", R. Penn and A. skjeltorp, Eds., New York , 1993.

9-Scheper, H. E.; Van, B. J. and Bassingth, W.; "Four Methods to Estimate the Fractal Dimension from self- Affme signals", IEEE Eng. In med. And Bio, Vol. 11, No.2, P. 57-64, June 1998.

10- Kountchev R., Milanova M., Todorov V., Kountcheva R., " Adaptive fuzzy filter for the reduction of blocking artifacts in images compressed with IDP decomposition and JPEG", WSEAS Transactions on Signal Processing, Vol .2, No .7, pp .941-8, 2006.

11- Siew, W.; chec, W. and chien, S., "An evolving Instructional Design Model for Web-based course: A case studying Malaysia", <http://www/swhng/mmu. edu. my>, December, 2009.

12- Frank Y .Shih, Yi-Tawu "Enhancement of image watermark retrieval based on genetic algorithms" Computer Vision Laboratory, College of Computing Sciences, New Jersey Institute of Technology, Newark, NJ 07102, United States Received 10 December 2002; accepted 18 May 2004 Available online 11 November 2004.