European Journal of Scientific Research
ISSN 1450-216X Vol.62 No.3 (2011), pp. 321-330

http://www.eurojournals.com/ejsr.htm

Performance Analysis of Clustering Algorithms in Brain Tumor Detection of MR Images

P. Tamije Selvy
Department of Computer Science & Engineering
V.L.B. Janakiammal College of Engineering & Technology
Coimbatore, India

V. Palanisamy
Info Institute of Engineering, Coimbatore, India
Department of Computer Science and IT

T. Purusothaman
Government College of Technology
Coimbatore, India

Abstract

The brain is the anterior most part of the central nervous system. Along with the Spinal cord, it forms the Central Nervous System (CNS). Brain tumor is an abnormal growth caused by cells reproducing themselves in an uncontrolled manner. Magnetic Resonance Imager (MRI) is the commonly used device for diagnosis. In MR images, the amount of data is too much for manual interpretation and analysis. Segmentation is an important process in most Medical Image Analysis. Clustering to magnetic resonance (MR) brain tumors maintains efficiency. Clustering is suitable for biomedical image segmentation as it uses unsupervised learning. This Paper analyses various clustering techniques to track tumor objects in Magnetic Resonance (MR) brain images. The input to this system is the MR image of the axial view of the human brain. The Clustering algorithms used are K-means, SOM, Hierarchical Clustering and Fuzzy C-Means Clustering. The given gray-level MR image is converted to a color space image and clustering algorithms are applied. The position of tumor objects is isolated from an MR image by using clustering algorithms The above clustering algorithms are analyzed and the performance is evaluated based on execution time and accuracy of the algorithms.

Keywords: Clustering, K-means, SOM, FCM, Hierarchical Clustering, Histogram Clustering.

1. Introduction

The brain is the anterior most part of the central nervous system.. Along with the spinal cord, it forms the Central Nervous System (CNS). The Cranium, a bony box in the skull protects it. The structure and function of the brain can be studied noninvasively by doctors and researchers using Magnetic Resonance Imaging (MRI).Magnetic Resonance Imaging (MRI), strongly depends on computer technology to generate or display digital images. Segmentation is an important process in most medical image analysis. It is very difficult to conduct surgery without using image processing techniques.

Complex medical processes cannot be done without image processing techniques. Structures like tumor, brain tissue and skull cannot be identified without image segmentation.

Image Segmentation[7][1][6] is needed to extract complex information from images. It takes a long time for diagnosis without using image processing techniques.

Clustering to Magnetic Resonance (MR) brain tumors maintains efficiency. Clustering is suitable for biomedical image segmentation as the number of clusters is usually known for images of particular regions of the human anatomy. This system uses color-based segmentation method. This system analyses various clustering techniques to track tumor objects in Magnetic Resonance (MR) brain images. The Clustering algorithms used are K-means, SOM, Hierarchical Clustering and Fuzzy C-Means Clustering.

A given gray-level MR image is converted into a color space image and clustering algorithms are applied. The position of tumor objects is separated from other items of an MR image by using clustering algorithms and histogram-clustering. In this system we combine, various clustering algorithms one by one and apply Histogram Clustering. After the clustering process, the cluster containing the tumor is selected as the primary segment. To eliminate the pixels which are not related to the tumor pixels, Histogram clustering is applied.

The performance analysis is conducted by taking a MRI Brain Tumor image as the input and applying all the four clustering algorithms to the image. The performance of the above four clustering algorithms are found based on the execution time and the number of tumor pixels.

2. MRI Segmentation Approaches

Magnetic Resonance Imaging (MRI) is an advanced medical imaging technique providing rich information about the human soft tissue anatomy. The goal of Magnetic Resonance (MR) image segmentation[6][7] is to accurately identify the principal tissue structures in these image volumes. However in MRI images, the amount of data is far too much for manual interpretation and analysis and this has been one of the biggest problems in the effective use of MRI. In the specific case of brain MRI, the problem of segmentation[1][12][13] is particularly critical for diagnosis and treatment purposes.

It is necessary to develop algorithms to obtain robust image segmentation such that the following may be observed:

- Automatic and semi-automatic delineation of areas to be treated to radio surgery.
- Delineation of tumors before and after surgical or radio-surgical intervention.
- Tissue classification: Volumes of white matter (WM), Grey matter (GM), Cerebrospinal fluid (CSF), Skull, Scalp and abnormal tissues.

The methods for MRI segmentation are described below.

Threshold Techniques

The classification of each pixel depends on intensity and color information. These techniques are efficient when the histograms of objects and background are clearly separated.

Edge-Based Methods

This method focuses on detecting contour. They fail when the image is too complex to identify a given border.

Region Based Segmentation

The concept of extracting features from a pixel and its neighbors is exploited to derive relevant information for each pixel.

Cooperative Hierarchical Computation Approach

This approach uses pyramid structures to associate the image properties to an array of father nodes.

Statistical Approaches

This approach labels pixels according to probability values which are determined based on the intensity distribution of the image.

3. Stages of the System

The system is logically separated into six stages for the ease of program design.

- Pseudo Color Translation.
- Color Space Translation
- Implementation of Clustering Algorithms.
- Cluster Selection
- Histogram Clustering
- Region Elimination

3.1. Pseudo Color Translation

Original MR Brain image is a gray-level image insufficient to support fine features. To obtain more useful features and enhance the visual density, the proposed method applies pseudo-color transformation, a mapping function that maps a gray-level pixel to a color-level pixel by a lookup table in a predefined color map. An RGB color map contains R, G, and B values for each item. Each gray value maps to an RGB item. The proposed method has adopted the standard RGB color map, which gradually maps gray-level values 0 to 255 into blue-to-green-to-red color.

3.2. Color Space Translation

To retrieve important features to benefit the clustering process, the RGB color space is further converted to a CIELab color model (L*a*b*). The L*a*b* space consists of a luminosity layer L*, a chromaticity-layer a*, which indicates where color falls along the red-green axis, and a chromaticity-layer b*, which indicates where the color falls along the blue-yellow axis.

3.3. Implementation of Clustering Algorithms

This system is implemented using four Clustering algorithms. They are,

- K-means Clustering Segmentation
- Clustering using Self Organizing Maps
- Hierarchical Clustering
- Fuzzy C-Means Clustering

3.3.1. K-means Clustering Segmentation

Algorithm: K-means

Step 1: The initial partitions are chosen by getting the R, G, B values of the pixels.

Step 2: Every pixel in the input image is compared against the initial partitions using the Euclidian Distance and the nearest partition is chosen and recorded.

Step 3: Then, the mean in terms of RGB color of all pixels within a given partition is determined. This mean is then used as the new value for the given partition.

Step 4: Once the new partition values have been determined, the algorithm returns to assigning each pixel to the nearest partition.

Step 5: The algorithm continues until pixels are no longer changing which partition they are associated with or until none of the partition values changes by more than a set small amount.

3.3.2. Clustering using Self Organizing Maps

Self-organizing maps (SOMs)[4][5] are data visualization techniques invented by Professor Teuvo Kohonen[4] which reduces the dimensions of data through the use of self-organizing neural networks. The problem that data visualization attempts to solve is that humans simply cannot visualize high dimensional data as is so techniques are created to help us understand this high dimensional data. The way SOMs go about reducing dimensions is by producing a map of usually 1 or 2 dimensions which plot the similarities of the data by grouping similar data items together. So SOMs accomplish two things, they reduce dimensions and display similarities.

Algorithm: SOM

Step 1: Randomize the map's nodes weight vectors.

Step 2: Grab an input vector.

Step 3: Traverse each node in the map

Step 4: Use Euclidean distance formula to find the similarity between the input vector and the map's node's weight vector.

Step 5: Track the node that produces the smallest distance (this node is the best matching unit, BMU)

Step 6: Update the nodes in the neighborhood of BMU by pulling them closer to the input vector.

$$Wv(t+1)=Wv(t)+ alpha(D(t)-Wv(t)) \quad (1)$$

Increment t and repeat from 2.

alpha-> monotonically decreasing learning coefficient. It is 1 for neurons close to BMU and zero for others.

D (t) -> input vector

Neighborhood function shrinks with time. At the beginning, when the neighborhood is broad, the self organizing takes place on a global scale. When the neighborhood has shrunk to just a couple of neurons, the weights are converging to local estimates.

3.3.3. Hierarchical Clustering

A Hierarchical clustering[5] method works by grouping data objects into a tree of clusters. This follows top down strategy. This algorithm starts with all objects in one cluster. It subdivides the cluster into smaller and smaller pieces until each object forms a cluster on its own or until it satisfies certain termination conditions such as desired number of clusters is obtained. This is called Divisive Hierarchical clustering.

Algorithm: Divisive Hierarchical Clustering

Step 1: The whole image is in one cluster.

Step 2: Find the most dissimilar point in the image and divide the image into two clusters.

Step 3: Repeat step2 for each cluster.

Step 4: A tree like structure is formed. Repeat step 2 until level 4 is reached. Level 4 has 8 clusters.

3.3.4. Fuzzy C-Means Clustering

The Fuzzy C-Means algorithm (often abbreviated to FCM) is an iterative algorithm[1] that finds clusters in data and which uses the concept of fuzzy membership. Instead of assigning a pixel to a single cluster, each pixel will have different membership values on each cluster. The Fuzzy C-Means[2] attempts to find clusters in the data by minimizing an objective function shown in the equation below:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} \mu_{ij}^{m} |x_i - c_j|^2 \tag{2}$$

J is the objective function. After one iteration of the algorithm the value of J is smaller than before. It means the algorithm is converging or getting closer to a good separation of pixels into clusters. N is the number of pixels in the image, C is the number of clusters used in the algorithm, and must be decided before execution, μ is the membership table -- a table of NxC entries which contains the membership values of each data point and each cluster, m is a fuzziness factor (a value larger than 1), xi is the i^{th} pixel in N, cj is j^{th} cluster in C and |xi - cj| is the Euclidean distance between xi and cj.

Algorithm: FCM

The input to the algorithm is the N pixels on the image and m, the fuzziness value. The fuzziness value of 2 is used in this system.

Step 1: Initialize μ with random values between zero and one; but with the sum of all fuzzy membership table elements for a particular pixel being equal to 1 -- in other words, the sum of the memberships of a pixel for all clusters must be one.

Step 2: Calculate an initial value for J using

$$J = \sum_{i=1}^{N} \sum_{j=1}^{c} \mu_{ij}^{m} |x_i - y_i|^2 \tag{3}$$

Step 3: Calculate the centroids of the clusters cj using,

$$c_j = \frac{\sum_{i=1}^{N} \mu_{ij}^{m} x_i}{\sum_{i=1}^{N} \mu_{ij}^{m}} \tag{4}$$

Step 4: Calculate the fuzzy membership table using

$$\mu_{ik} = \frac{1}{\sum_{k=1}^{c} \left(\frac{|x_i - c_j|}{|x_i - c_k|} \right)^{\frac{2}{m-1}}} \tag{5}$$

Step 5: Recalculate J.

Step 6: Go to step 3 until a stopping condition was reached.

Some possible stopping conditions of the algorithm are:

1. When a number of iterations were executed, we can consider that the algorithm achieved a "good enough" clustering of the data.
2. The difference between the values of J in consecutive iterations is small (smaller than a user-specified parameter ε), therefore the algorithm has converged.

Defuzzification:

1. At the end of the execution of the algorithm we have, for each pixel, the membership values for that pixel in each cluster.
2. Traditionally the algorithm can then defuzzify its results by choosing a "winning" cluster, i.e. the one which is closer to the pixel in the feature space, is the one for which the membership value is highest; and using that cluster center as the new values for the pixel.

3.4. Cluster Selection

After the clustering process, the cluster containing an area of interest (tumor) is selected as the primary segment.

3.5. Histogram Based Clustering

To eliminate the pixels which are not related to the interest in the selected cluster, histogram clustering is applied by luminosity feature L* and color information a* and b* to derive the final segmented result. The K-means algorithm uses L* for histogram clustering where as SOM uses a* and b*. The histogram clustering in hierarchical segmentation uses l* to achieve the final segmentation result. In RGB color space, histogram clustering uses red value to derive the final segmentation result.

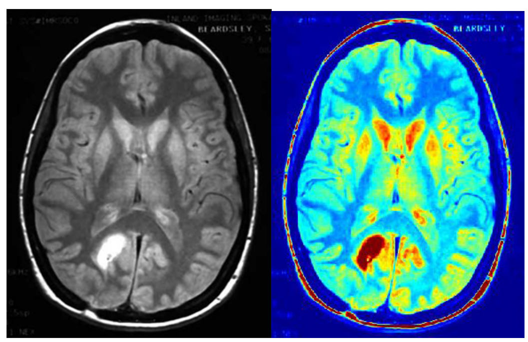
3.6. Region Elimination

The output of histogram clustering consists of tumor region as well as the other regions which has the same luminance and color values as the tumor. The regions which are smaller than the tumor are eliminated using region growing algorithm.

4. Experimentation and Results

In the proposed system, brain tumors were segmented by using four clustering algorithms. The results obtained for the four clustering algorithms are given in this section. Figure 1 shows the original MRI brain tumor image as well as the image obtained after pseudo color translation.

Figure 1: (a) Original MRI brain tumor image (b) Colored MRI image

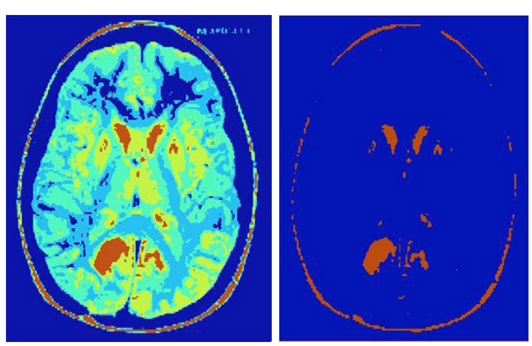


(a) (b)

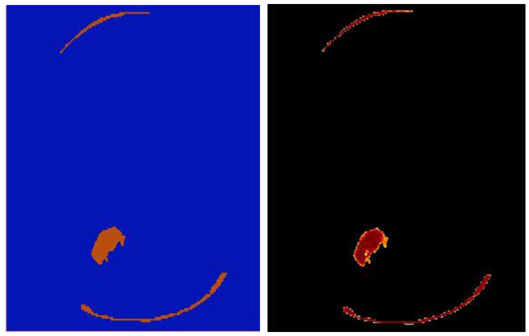
4.1. K-means Clustering Algorithm

Figure 1 shows the original MRI brain tumor image. And the image after Pseudo color translation. Figure 2 shows the image after K-means Clustering in L*a*b* color space. The number of clusters is given by the user as 5. Figure 2 shows the image obtained after K-means Clustering Segmentation, Cluster selection, region elimination and the segmented tumor image respectively.

Figure 2: (a) K-means clustered image (b) Image after Cluster Selection (c) Image after region elimination (d) Segmented Tumor image



(a) (b)

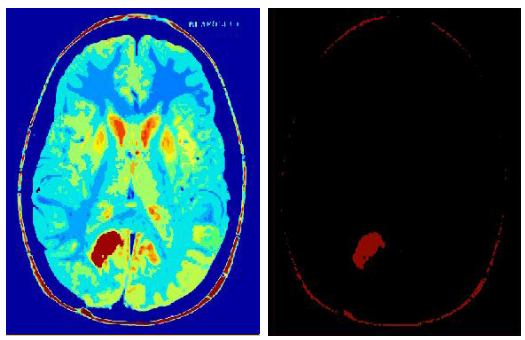


(c) (d)

4.2. Self Organizing Map

The input is a MRI brain tumor image after Pseudo Color Translation as shown in Figure 1. Figure 3 shows the image after Clustering using Self Organizing Map neural network. The clustering is in L*a*b* color space. The number of iterations is given by the user as 250. Figure 3 shows the image after cluster selection and histogram clustering.

Figure 3: (a) SOM clustered image (b) Segmented tumor after histogram clustering.

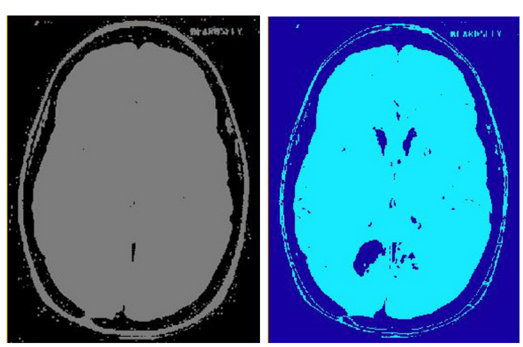


(a) (b)

4.3. Hierarchical Clustering

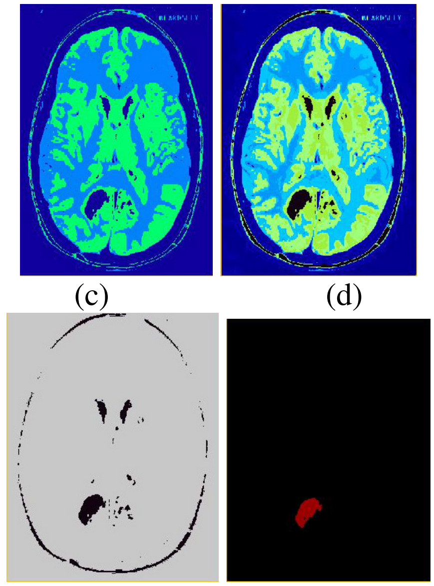
The input is a MRI brain tumor image after Pseudo Color Translation as shown in Figure 1. Figure 4 shows the level1 image after Hierarchical Clustering[1] which has only one cluster, the level2 image which has two clusters, the level3 image which has four clusters, and the level4 image which has eight clusters, the image after cluster selection and after region elimination respectively. The clustering is in RGB color space.

Figure 4: (a) Level1 clustered image (b) Level2 clustered image (c) Level3 clustered image (d) Level4 clustered image (e) Image after histogram clustering (f) Segmented tumor image.



(a) (b)

Figure 4: (a) Level1 clustered image (b) Level2 clustered image (c) Level3 clustered image (d) Level4 clustered image (e) Image after histogram clustering (f) Segmented tumor image. - continued



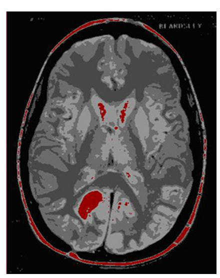
(c) (d)

(e) (f)

4.4. Fuzzy C-Means Clustering

The input is a MRI brain tumor image after Pseudo Color Translation as shown in Figure 1. Figure 5 shows the image after Clustering using Fuzzy C-Means Clustering [1] Algorithm. The clustering is in L*a*b* color space. The number of iterations is given by the user as 100. The number of clusters is given by the user as 5.

Figure 5: (a) FCM clustered image



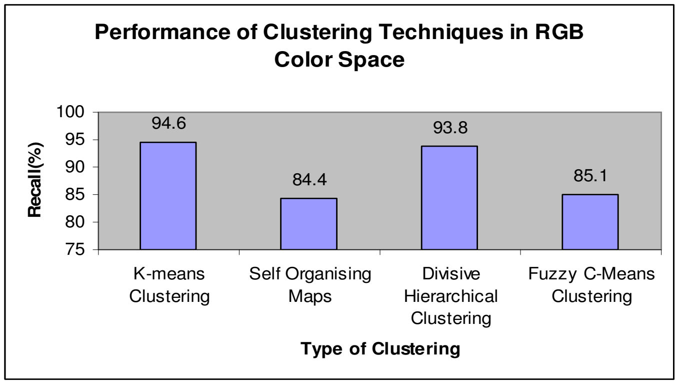
(a)

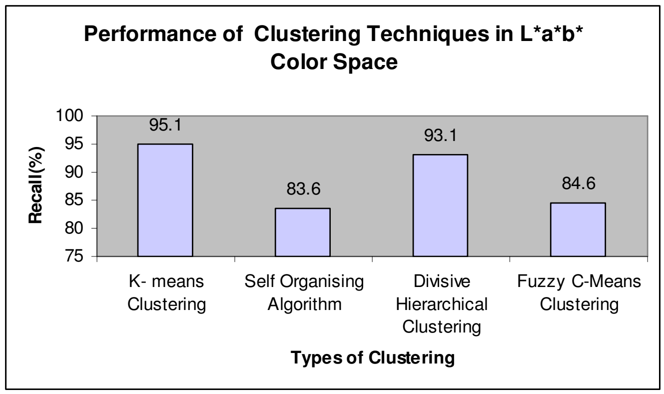
Table 1: Performance of Clustering Techniques in RGB Color Space

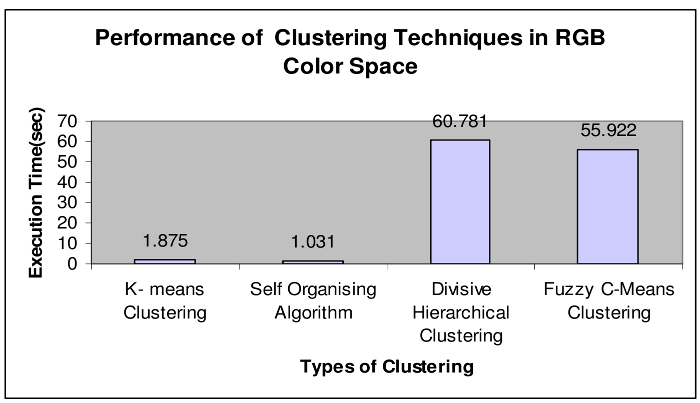
Type of Clustering	Recall	Execution Time (sec)
K-means Clustering	94.6	1.875
Self Organising Maps	84.4	1.031
Divisive Hierarchical Clustering	93.8	60.781
Fuzzy C-Means Clustering	85.1	55. 922

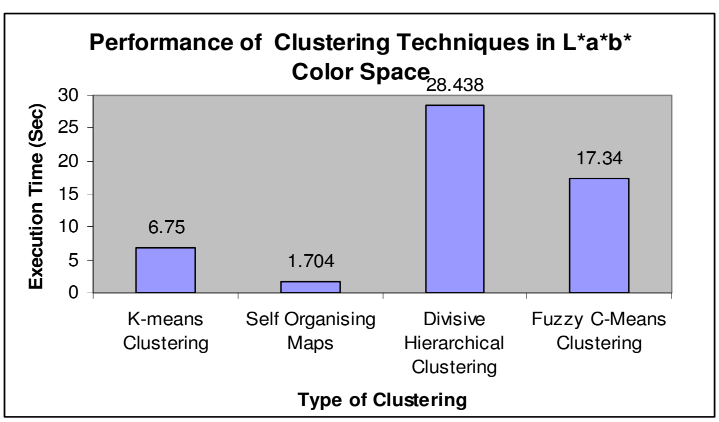
Table 2: Performance of Clustering Techniques in L*a*b* Color Space

Type of Clustering	Recall	Execution Time (sec)
K-means Clustering	95.1	6.75
Self Organising Maps	83.6	1.704
Divisive Hierarchical Clustering	93.1	28.438
Fuzzy C-Means Clustering	84.6	17.340


Performance of Clustering Techniques in RGB Color Space
Recall(%)
100
95
90
85
80
75
94.6
84.4
93.8
85.1
K-means Clustering
Self Organising Maps
Divisive Hierarchical Clustering
Fuzzy C-Means Clustering
Type of Clustering


Performance of Clustering Techniques in L*a*b* Color Space
Recall(%)
100
95
90
85
80
75
95.1
83.6
93.1
84.6
K- means Clustering
Self Organising Algorithm
Divisive Hierarchical Clustering
Fuzzy C-Means Clustering
Types of Clustering


Performance of Clustering Techniques in RGB Color Space
Execution Time(sec)
70
60
50
40
30
20
10
0
1.875
1.031
60.781
55.922
K- means Clustering
Self Organising Algorithm
Divisive Hierarchical Clustering
Fuzzy C-Means Clustering
Types of Clustering


Performance of Clustering Techniques in L*a*b* Color Space
Execution Time (Sec)
30
25
20
15
10
5
0
6.75
1.704
28.438
17.34
K-means Clustering
Self Organising Maps
Divisive Hierarchical Clustering
Fuzzy C-Means Clustering
Type of Clustering

5. Conclusion and Future Enhancements

In this system brain tumors have been segmented with the help of four methods. The execution time for K-means Clustering and SOM were less compared to the other clustering methods. Regarding the number of tumor pixels, K-means clustering and Hierarchical clustering gave a better result than the other methods. The four clustering algorithms were tested with a database of 100 MRI brain images. K-means and Hierarchical clustering achieved about 95% result. SOM and FCM achieved a result of about 80%. In this system, the axial view of the human brain is taken for tumor detection. The system can be extended to detect tumors on in other views of the brain. This system considers the color and luminosity parameters for tumor detection. The texture can be taken as an additional parameter for tumor detection

References

[1] S.R. Kannan , Jan. 2005 “Segmentation of MRI using new Unsupervised Fuzzy C means algorithm” ICGST-GVIP Journal, Volume 5, Issue2.

[2] Juraj Horvath, 2006 “Image Segmentation using Fuzzy C-means” SAMI 2006.

[3] Lei jiang,Wenhui Yang , 2003 “A modified Fuzzy C-means algorithm for Magnetic Resonance Images” Proc. VIIth Digital Image Computing: Techniques and Applications,

[4] Constantino, Reyes,Aldasoro, “ Image Segmentation with Kohonen Neural Network SOM” CIIST, 2004

[5] Christophe Ambroise,Fouad Badrean,Sylvie Thira,Genive Seze, Hierarchical Clustering of SOM for Cloud Classification, NeuroComputing, Volume 30, January 2000, Pages 47-52

[6] Fu, K.S. and J.K. Mui, 1981.”A survey on image segmentation”, Pattern Recognition,13: 3-16.

[7] Gonzalez, R.C. and R.E. Woods, “Digital image processing, Pearson Education”, 2002.

[8] Haralick, R.M. and L.G. Shapiro, 1985. “Survey ,image segmentation techniques”, Computer Vision, Graphics Image Process, 29: 100-132[9]

[9] J. Goldberger, S. Gordon, H. Greenspan, “Unsupervised Image-Set Clustering Using an Information Theoretic Framework”, IEEE Transactions on Image Processing, Vol. 15, No. 2, February 2006, pp 449-458.

[10] N. Pal and S. Pal, "A review on image segmentation techniques," Pattern Recognition, vol. 26, pp. 1277-1294, 1993.

[11] Y. Zhang, "A survey on evaluation methods for image segmentation," Pattern Recognition, vol. 29, pp. 1335-1346, 1996.

[12] A. Worth, N. Makris, V. Caviness, and D. Kennedy, "Neuroanatomical segmentation in MRI: Technological objectives," International Journal of Pattern Recognition and Artificial Intelligence, vol. 11, pp. 1161-1187, 1997.

[13] M. Schmidt, I. Levner, R. Greiner, A. Murtha, and A. Bistritz. "Segmenting Brain Tumors using Alignment-Based Features". International Conference on Machine Learning and Applications, Los Angeles, December 2005.

[14] C. Majós, M. Julià-Sapé, J. Alonso, M. Serrallonga, C. Aguilera, J. Acebes, C. Arús and J. Gili. “Brain Tumor Classification by Proton MR Spectroscopy: Comparison of Diagnostic Accuracy at Short and Long TE”. American Journal of Neuroradiology 25:1696-1704, November-December 2004.