



Optimized K-Means and Fuzzy C-Means for MRI Brain Image Segmentation

A Pethalakshmi

*HOD & Associate Professor
Department of Computer Science
MVM Government Arts College, Dindigul
Tamilnadu, India
pethalakshmi@yahoo.com*

A Banumathi

*Assistant Professor
Department of Computer Science
Government Arts College, Karur
Tamilnadu, India
banukarthikeyan7811@gmail.com*

Abstract- Image segmentation is an important aspect of medical image processing, where K-Means and Fuzzy C-Means clustering approaches are widely used in biomedical applications particularly for brain tumor detection. In this paper, K-Means and Fuzzy C-Means clustering algorithms are analyzed and found it have two major drawbacks. First drawback, it forces the objects to be clustered within the user defined K number of clusters. Without having prior knowledge on the database it is difficult to predict the number of resultant cluster that is to be obtained. Second drawback, the quality of the resultant cluster is based on the initial seeds where it is selected randomly. On random selection there is possibility of selecting nearby seeds as a centroid for each cluster of K. In this case also once again the algorithm forces all the data's towards the fixed centroid of each cluster and hence there is possibility of wrong diagnosing happens. But, the brain is highly sensitive and control centered organ because of above mentioned drawbacks, on implementing these algorithms may leads to wrong diagnosing which is a life risk job. In order to overcome this drawback the current paper focused on developing the Unique Clustering with Affinity Measure (UCAM) and Fuzzy-UCAM algorithm for clustering without defining initial seed and number of resultant clusters. Unique clustering is obtained with the help of affinity measures.

Keywords- K-Means, fuzzy C-Means, UCAM, fuzzy-UCAM, clusters, image segmentation, MRI

I. INTRODUCTION

Biomedical analysis has evolved considerably over the last couple of decades. The widespread availability of suitable detectors has aided the rapid development of new technologies for the monitoring and diagnosis, as well as treatment, of patients. Over the last century technology has advanced from the discovery of x-rays to a variety of imaging tools such as MRI, Computed ultrasonography [2]. Three-dimensional (3-D) processing and Tomography (CT), Positron Emission Tomography (PET). MRI has provided a means for imaging tissue at very high resolutions providing the desired information for use in fields like reparative surgery, radiotherapy treatment planning, stereotactic neurosurgery, and others [6,7].

Image segmentation plays a major role in the field of biomedical applications. The segmentation technique is widely used by the radiologists to segment the input medical image into meaningful regions [1,4,5,8]. The specific application of this technique is to detect the tumor region by segmenting the abnormal MR input image. The slice of the tumor region can be tracked using techniques like K-Means and fuzzy C-Means, which aid the radiologists in treatment planning. The primitive techniques are based on manual segmentation is a time consuming process besides being susceptible to human errors. Several automated techniques have developed which the drawbacks of manual segmentation.

Clustering is one of the widely used image segmentation techniques which classify patterns in such a way that samples of the same group are more similar to one another than samples belonging to different groups [3,9,10]. The K-Means algorithm is a partition clustering method that separates data into K groups. Main drawback of this algorithm is that of a priori fixation of number of clusters and seeds. There has been considerable interest recently in the use of fuzzy clustering methods, which retain more information from the original image than hard clustering methods. Fuzzy C-Means algorithm is widely preferred because of its additional flexibility which allows pixels to belong to multiple classes with varying degrees of membership. But the major operational complaint is that the FCM technique is time consuming. The drawback of both K-Means and fuzzy C-Means is improved by UCAM and Fuzzy-UCAM algorithm.

II. RELATED WORK

A. Image segmentation

Segmentation subdivides an image into its constituent regions or objects [12]. The level of detail to which the subdivision is carried depends on the problem being solved. That is, segmentation should stop when the objects or regions of interest in an application have been detected. Segmentation of nontrivial images is one of the most difficult tasks in image processing. Segmentation accuracy determines the eventual success or failure of computerized analysis procedures. For this reason, considerable care should be taken to improve the probability of accurate segmentation.

Much past work on medical image segmentation relied strictly on human graphical interaction to define regions, using methods such as manual slice editing, region painting and interactive thresholding. Rajapakse [13] classified the different methods of image segmentation as four main categories. (1) The classical methods such as thresholding region growing and edge based techniques. (2) The statistical methods such as the maximum-likelihood-classifier (MLC). These methods are basically supervised and depend on the prior model and its parameters. Vannier et al. [14] reported satisfactory preliminary results with Bayesian MLC. Ozkan [15] made a comparison between the MLC and the neural network classifier which showed the superiority of the neural network. New methods of segmentation that could be classified as statistical methods have been introduced in the past few years. Hansen [16] used a probabilistic supervised relaxation technique for segmenting 3D medical images. The method introduced the use of cues to guide the segmentation. Those cues marked by the user have the mean and standard deviation as description parameters. (3) The neural networks methods one example of which is the work of Ahmed et al. [17] who used a two stages neural network system for CT/MRI image segmentation. The first stage is a self-organized principal component analysis (SOPCA) network and the second stage consists of a self-organizing feature map (SOFM). The results obtained compare favorably with the classical and statistical methods. (4) The Fuzzy Clustering methods. In [18] a comparison between the fuzzy clustering and neural network techniques in segmenting magnetic resonance images of the brain debated for the need of unsupervised technique in segmentation which was provided using the unsupervised fuzzy c- mean algorithm. However the long time taken by the fuzzy C-mean algorithm was documented. And for initial clustering it uses K-Means clustering method where it has the drawback of prior fixation of seeds and cluster K. The complaints that are found in K-Means and fuzzy C-Means have been rectified through UCAM and Fuzzy-UCAM which is the newly proposed in this paper.

B. MRI

Magnetic Resonance Imaging (MRI) is an imaging technique used primarily in medical settings to produce high quality images of the inside of the human body. In order to preview about MRI, in this section we give a brief description of the principles of MRI which are referred to [11]. In MRI, the image is a map of the local transverse magnetization of the hydrogen nuclei. This transverse magnetization in turn depends on several intrinsic properties of the tissue. The MR phenomenon relies on the fundamental property that protons and neutrons that make up a nucleus possess an intrinsic angular momentum called spin. When protons and neutrons combine to form nucleus, they combine with oppositely oriented spins. Thus, nuclei with an even number of protons and neutrons have no net spin, whereas nuclei with an odd number of protons or neutrons possess a net spin. Hydrogen nuclei have an NMR signal since its nucleus is made up of only a single proton and possess a net spin. The human body is primarily made up of fat and water, which have many hydrogen atoms. Medical MRI primarily images the MRI signal from the hydrogen nuclei in the body tissues.

The net spin of the nucleus around its axis gives it an angular moment. Since the proton is a positive charge, a current loop perpendicular to the rotation axis is also created, and as a result the proton generates a magnetic field. The joint effect of the angular moment and the self generated magnetic field gives the proton a magnetic dipole moment parallel to the rotation axis. Under normal condition, one will not experience any net magnetic field from the volume since the magnetic dipole moments are oriented randomly and on average equalize one another. When placed in a magnetic field, a proton with its magnetic dipole moment precesses around the field axis. The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document.

The frequency of this precession, ν_0 , is the resonant frequency of NMR and is called the Larmor frequency. The precession frequency is directly proportional to the strength of the magnetic field, i.e.

$$\nu_0 = gB_0 \quad (1)$$

Where B_0 is the main magnetic field strength, and g is a constant called gyromagnetic ratio which is different for each nucleus (42.56 MHz/Tesla for protons). Given a specimen, the application of a magnetic field

B_0 would create a net equilibrium magnetization M_0 per cubic centimeter, which is aligned to the B_0 field. The M_0 is the net result of summing up the magnetic fields due to each of the H nuclei and is directly proportional to the local proton density (or spin density). However, M_0 is many orders of magnitude weaker than B_0 and is not directly observable. By tipping M_0 away from the B_0 field axis with an appropriate RF pulse having a frequency equals to the Larmor frequency, a longitudinal magnetization component ML and a transverse magnetization component MT is produced. When the RF pulse is turned off, the longitudinal magnetization component ML recovers to M_0 with a relaxation time $T1$, and the transverse magnetization component MT dephases and decays to zero with a relaxation time $T2$.

During relaxation, the protons lose energy by emitting their own RF signal with the amplitude proportional to MT . This signal is referred to as the *free-induction decay* (FID) response signal. $T2$ indicates the time constant required for the FID response signal from a given tissue type to decay. The FID response signal is measured by an RF coil placed around the object being imaged. When MR images are acquired, the RF pulse is repeated at a predetermined rate. The period of the RF pulse sequence is the *repetition time*, TR . The FID response signals can be measured at various times within the TR interval. The time between which the RF pulse is applied and the response signal is measured is the *echo delay time*, TE . The TE is the time when the spin echo occurs due to the refocusing effects of the 180 degree refocusing pulse applied after a delay of $TE/2$ from the RF pulse. The TR and TE control how strongly the local tissue relaxation times, $T1$ and $T2$, affect the signal. By adjusting TR and TE the acquired MR image can be made to contrast different tissue types and provide clear quality images. On processing these quality images for segmentation through the techniques like clustering, this gives silver lining for tumor deduction.

C. K-Means

In recent trends of biomedical field, clustering is one of the most powerful techniques for deducting tumor in human brain MR images. The main objective in cluster analysis is to group object that are similar in one cluster and separate objects that are dissimilar by assigning them to different clusters. One of the most popular clustering methods is K-Means clusters algorithm. It classifies object to pre-defined number of clusters, which is given by the user (assume K clusters). The idea is to choose random cluster centers, one for each cluster. These centers are preferred to be as far as possible from each other. In this algorithm Euclidean distance measure is used between two multidimensional data points.

$$X = (x_1, x_2, x_3, \dots, x_m) \tag{2}$$

$$Y = (y_1, y_2, y_3, \dots, y_m) \tag{3}$$

The Euclidean distance measure between the above points x and y are described as follows:

$$D(X, Y) = (\sum (x_i - y_i)^2)^{1/2} \tag{4}$$

The K-Means method aims to minimize the sum of squared distances between all points and the cluster centre. This procedure consists of the following steps, as described below

Algorithm1: K-Means clustering algorithm

Input: $D = \{d_1, d_2, d_3, \dots, d_n\}$ // Set of n data points.

K - Number of desired clusters

Output: A set of K clusters.

Steps:

1. Select the number of clusters. Let this number be K.
2. Pick K seeds as centroids of the K clusters. The seeds may be picked randomly unless the user has some insight into the data.
3. Compute the Euclidean distance of each object in the dataset from each of the centroids.
4. Allocate each object to the cluster it is nearest to based on the distances computed in the previous step.
5. Compute the centroids of the clusters by computing the means of the attribute values if the objects in each cluster.
6. Check if the stopping criterion has been met (e.g. the cluster membership is unchanged). If yes, go to step 7. If not go to step 3.

7. [Optional] One may decide to stop at this stage or to split a cluster or combine two clusters heuristically until a stopping criterion is met.

Though the K-Means algorithm is simple, but it has some drawbacks in its quality of the final clustering, since it is highly depends on the initial centroids and hence on implementing this technique may mislead the direction on deducting of tumor. Because of the random selection of seeds in K-Means, sometimes it has the possibility of cluster error when it selects nearby seeds as centroid for K cluster. Even though MRI is quite sufficient for deducting the tumor, due to the drawbacks in this technique it makes the total process as inefficient. To improve the quality of the cluster and to provide unique clustering above drawback is rectified through newly proposed UCAM algorithm as explained in the research methodology section.

D. Fuzzy C-Means

The fuzzy C-Means clustering algorithm is one of the popular traditional techniques in biomedical field for deducting tumor. The fuzzy C-Means [19] is a variation of the popular K-Means clustering algorithm, in which a degree of membership of clusters is incorporated for each data point. The centroids of the clusters are computed based on the degree of memberships as well as data points. The random initialization of memberships of instances used in both traditional fuzzy C-Means and K-Means algorithms lead to the inability to produce consistent clustering results and often result in undesirable clustering results[9]. This algorithm works by assigning membership to each data point corresponding to each cluster center on the basis of distance between the cluster center and the data point. Clearly, summation of membership of each data point should be equal to one. After each iteration membership and cluster centers are updated according to the formula.

One of the most widely used fuzzy clustering algorithms is the Fuzzy C-Means (FCM) Algorithm (Bezdek 1981). The FCM algorithm attempts to partition a finite collection of n elements $X = \{x_1, \dots, x_n\}$ into a collection of C fuzzy clusters with respect to some given criterion. Given a finite set of data, the algorithm returns a list of C cluster centres $C = \{c_1, \dots, c_c\}$ and a partition matrix $U = u_{ij} \in [0,1] i=1, \dots, n, j=1, \dots, C$ where each element u_{ij} tells the degree to which element x_i belongs to cluster C_j . Like the K-Means algorithm, the FCM aims to minimize an objective function. The standard function is

$$u_k(x) = \frac{1}{\sum_j \left(\frac{d(\text{center}_k, x)}{d(\text{center}_j, x)} \right)^{2/(m-1)}} \quad (5)$$

which differs from the K-Means objective function by the addition of the membership values u_{ij} and the fuzzifier m . The fuzzifier m determines the level of cluster fuzziness. A large m results in smaller memberships u_{ij} converge to 0 or 1, which implies a crisp partitioning. In the absence of experimentation or domain knowledge, m is commonly set to 2. The basic FCM Algorithm, given n data points (x_1, \dots, x_n) to be clustered, a number of c cluster with (C_1, \dots, C_C) the center of the clusters, and m the level of cluster fuzziness.

Any point x has a set of coefficients giving the degree of being in the k th cluster $w_k(x)$. With fuzzy C-Means, the centroid of a cluster is the mean of all points, weighted by their degree of belonging to the cluster:

$$C_k = \frac{\sum_x w_k(x) x}{\sum_x w_k(x)} \quad (6)$$

The degree of belonging $w_k(x)$, is related inversely to the distance from x to cluster center as calculated on the previous pass. It also depends on a parameter m that controls how much weight is given to the closest centre. The fuzzy C-Means algorithm is very similar to the K-Means algorithm.

Algorithm 2: Fuzzy C-Means clustering algorithm

Input: $D = \{d_1, d_2, d_3, \dots, d_n\}$ // Set of n data points.

C - Number of desired clusters

Output: A set of C clusters, degree of membership matrix

Steps:

1. Choose a C number of clusters.
2. Assign randomly to each point coefficients for being in the clusters.
3. Repeat until the algorithm has converged (that is, the coefficients' change between two iterations is no more than ϵ , the given sensitivity threshold)
4. Computer the centroid for each cluster, using the formula (6)
5. For each point, computes its coefficients of being in the clusters, using the formula (5).

The algorithm minimizes intra-cluster variance as well, but has the same problems as K-Means; that the results depend on the initial choice of k and its centroid, since it uses K-Means for its initial clustering. The major operational complaint is that the FCM technique is time consuming. And hence if this method is applied in MRI to segment for deducting tumor has the possibility of cluster error and which is life risk matter for the patients. Listed above two drawbacks were eradicated through another one newly proposed algorithm named as fuzzy-UCAM which purely works on affinity measure.

III. RESEARCH METHODOLOGY

A. The Proposed Method : 1

In cluster analysis, one does not know what classes or clusters exist and the problem to be solved is to group the given data into meaningful clusters. Here on the same motive UCAM algorithm is developed. UCAM algorithm is a clustering algorithm basically for numeric data. It mainly focuses on the drawback of K-Means clustering algorithm. In K-Means algorithm, the process is initiated with the initial seeds and number of cluster is obtained. But the number of cluster that is to be obtained cannot be predicted on a single view of the dataset. The result may not be unique if the number of cluster and the initial seed is not properly identified.

UCAM algorithm is implemented with the help of affinity measure for clustering. The process of clustering in UCAM initiated without any centorid and number of clusters that is to be produced. But it set the threshold value for making unique clusters. The step by step procedure for UCAM are given below

Algorithm 3: The UCAM algorithm

Input: $D = \{d_1, d_2, d_3, \dots, d_n\}$ // Set of n data points.

S – Threshold value.

Output: Clusters. Number of cluster depends on affinity measure.

UCAM Algorithm Steps:

1. Set the threshold value T.
2. Create new cluster structure if it is the first tuple of the dataset.
3. If it is not first tuple compute similarity measure with existing clusters.
4. Get the minimum value of computed similarity S.
5. Get the cluster index of C_i which corresponds to S.
6. If $S \leq T$, then add current tuple to C_i .
7. If $S > T$, create new cluster.
8. Continue the process until the last tuple of the dataset.

Uniqueness of the cluster is depends on the initial setting of the threshold value. If the threshold value increases number of cluster decreases. In UCAM there is no initial prediction on number of resultant cluster. Here, in this algorithm resultant cluster purely based on the affinity measure and hence it rectifies the major drawbacks of K-Means that the random selection of seeds and prior fixation of seeds. On applying UCAM clustering algorithm for tumor deduction it works efficiently by fixing the threshold value. By decreasing the threshold value it forms with more unique cluster and which helps in finding the exact area of tumor. And by increasing the threshold value method finds the tumor along with area affected by the tumor.

B. The Proposed Method : 2

The fuzzy-UCAM clustering algorithm is a variation of the UCAM clustering algorithm, in which a degree of membership of clusters is incorporated for each data point. The centroids of the clusters are computed based on the members of the cluster. The random initialization of the process of traditional fuzzy C-Means algorithms leads to cluster error and affects the uniqueness of the cluster. Fuzzy-UCAM algorithm works to rectify the cluster error and increase the uniqueness of Fuzzy C-Means through affinity measure. The Fuzzy-UCAM algorithm is outlined as follows

Algorithm 4: The Fuzzy-UCAM algorithm

Input: $D = \{d_1, d_2, d_3, \dots, d_n\}$ // Set of n data points.

S – Threshold value.

Output: Resultant Clusters, Degree of membership matrix.

Fuzzy-UCAM Algorithm Steps:

1. Set the threshold value T.
2. Create new cluster structure if it is the first tuple of the dataset.
3. If it is not first tuple compute similarity measure with existing clusters.
4. Get the minimum value of computed similarity S.
5. Get the cluster index of C_i which corresponds to S.
6. If $S \leq T$, then add current tuple to C_i .
7. If $S > T$, create new cluster.
8. Continue the process until the last tuple of the dataset.
9. Compute membership matrix for all data points in the resultant cluster using the formula (5).

On implementing Fuzzy-UCAM clustering algorithm, which produces the result with the $M \times N$ matrix, where M is number of resultant cluster and N is the total number of objects in the initial set which is to be clustered. The each row of the matrix indicates the degree of membership of the particular object towards all the clusters. The sum of each row should be the values between 0 and 1.

Fuzzy-UCAM algorithm results with unique clusters which are free from cluster error, since it is basically uses UCAM clustering algorithm result for its fuzzy measure. The number of resultant cluster is depends up on the threshold value, if the threshold value increases then the number of resultant clusters decreases and on decreasing the number of resultant cluster increases. On implementing fuzzy-UCAM in MRI for tumor deduction by adjusting the threshold value it not only finds exact area of tumor and areas affected by tumor but also give to what degree of representation it has been affected through membership matrix.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

UCAM and Fuzzy-UCAM clustering method is basically designed for numerical data and hence MR image is converted to numerical data for each and every pixel of image by using matlab commands. These numerical data is passed into proposed method as an input and the resultant cluster is also obtained in the form of numerical representation, which is once again converted to image to view the area of tumor in image form. Before analyzing tumor deduction section-a and b gives the clear view on the efficiency of UCAM and fuzzy-UCAM on a numerical data when it is applied in a small numerical data set. Section-c and d deals with tumor analysis and time consuming factor of existing and proposed method. Section-e represents the comparative analysis.

A. K-Means and UCAM clustering

The K-Means algorithm is implemented in a very small sample data with ten student's information. The process of K-Means clustering is initiated with three initial seeds, which results with three clusters as notated below

$$\begin{aligned} C_1 &= \{ S_1, S_9 \} \\ C_2 &= \{ S_2, S_5, S_6, S_{10} \} \\ C_3 &= \{ S_3, S_4, S_7, S_8 \} \end{aligned}$$

Where S_1, S_2, \dots, S_{10} Student's details which considers only numeric attributes. In the above study of K-Means clustering algorithm results with three clusters where low marks and high marks are found in all clusters, since the initial seeds do not have any seeds with the marks above 90. Hence if the initial seeds not defined properly then the result won't be unique and more over it has been constrained that it should have only three clusters. In K-Means the initial seeds are randomly selected and hence result of two executions on the same data set will not get the same result unless the initial seeds are same. The main drawback in K-Means is that initial seeds and number of cluster should be defined though it is difficult to predict it, in the early stage.

UCAM algorithm has also been the sample data performs for the K-Means. The process is initiated with threshold value T and results with following clusters as shown below

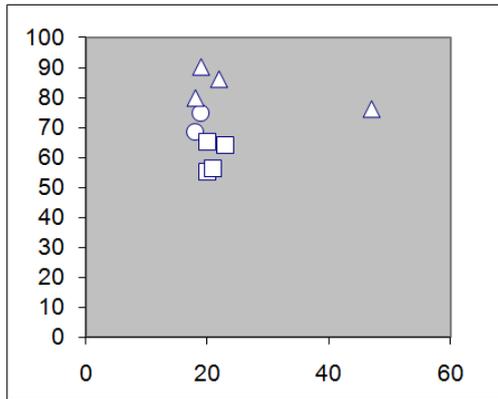
$$\begin{aligned} C_1 &\rightarrow \text{Cluster with medium marks.} \\ C_2 &\rightarrow \text{Cluster with high marks.} \\ C_3 &\rightarrow \text{Cluster with low marks.} \end{aligned}$$

$$C4 = \{ S_9 \}$$

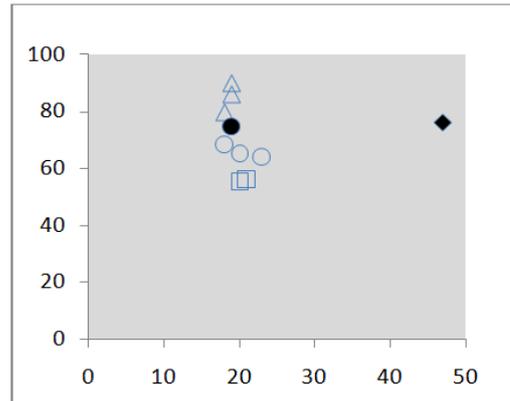
$$C5 = \{ S_{10} \}$$

S_9 and S_{10} are found to be having peculiar characteristics for the given threshold value. These two objects have major dissimilarity with the existing clusters and hence it cannot merge with other clusters. By increasing the threshold value it can be merged with other cluster based on the user requirements, but it reduces the cluster uniqueness and hence it proves that UCAM clustering algorithm has the flexibility of obtaining both approximate clustering and unique clustering.

The cluster representation of K-Mean and UCAM are illustrated through scatter graph as shown below in which each symbol indicates a separate cluster.



Figures 1 (a) : Clustering through K-Means

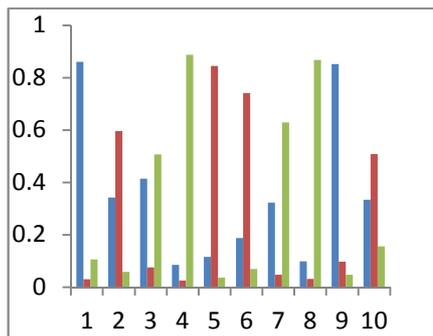


(b)Clustering through UCAM

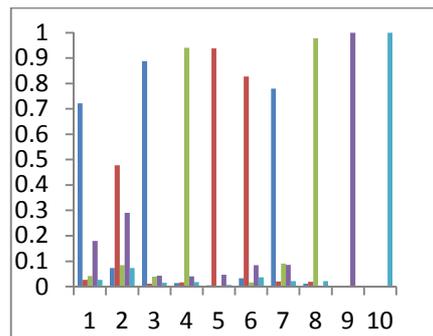
In the above graph each symbol represents a separate cluster. In Figure 1 (a) shows the clustering overlaps with each other but in Figure 1 (b) all the cluster are unique in representation compared to K-Means clustering and the dark shaded symbols are peculiar objects, based on the application it can be projected out otherwise it can be merged with nearby cluster by adjusting the threshold value. Both approximate clustering and unique cluster can be obtained by increasing and decreasing the threshold values.

B. Fuzzy C-Means and Fuzzy-UCAM

Uniqueness of the clusters for fuzzy C-Means and fuzzy-UCAM is measured by using the same data that were used in K-Means and UCAM. The membership matrix of fuzzy C-Means and fuzzy-UCAM is illustrated in the following bar chart representation



Figures 2 (a) : Clustering through Fuzzy C-Means



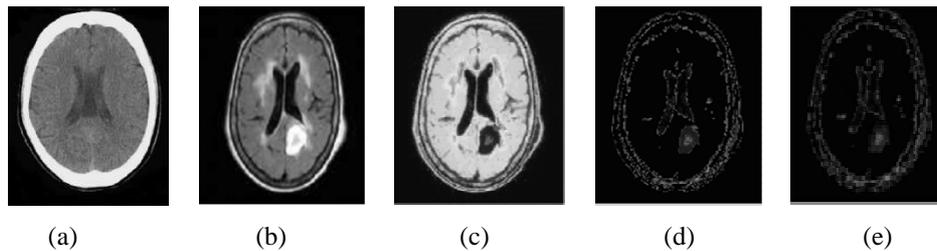
(b)Clustering through Fuzzy-UCAM

Figure.2 gives the clear visualization on cluster uniqueness of numerical data representation on Fuzzy C-Means and Fuzzy-UCAM. Each series indicate the possibility of particular data into all other possible clusters. If one object is classified into particular cluster then the degree of possibility towards other cluster is least significant in the case of clustering through Fuzzy-UCAM figure.2 (b). But in Fuzzy C-Means clustering it has the reasonable degree of possibility toward other clusters as shown in figure.2 (a). The above chart gives clear view on the higher degree of uniqueness in clustering by Fuzzy-UCAM compared to Fuzzy C-Means.

C. Analysis of the Area of Tumor

Radiologic diagnosis is based on the multiparametric imaging profile through CT, conventional MRI, advanced MRI [21]. Tumor characterization is difficult because neoplastic tissue is often heterogeneous in spatial and imaging profile, and for some imaging techniques often overlaps with normal tissue. Fuzzy-UCAM clustering algorithm is used to reduce overlaps with the normal tissue with tumor in segmentation of image for the analysis of the exact area affected with tumor. Since, the brain is highly sensitive and control centered organ in which exact diagnosing is must to avoid a life risk job.

The results obtained from the experimentation on the proposed segmentation approach are presented in the section. The proposed segmentation approach has been programmed in matlab. Before applying the proposed function, the image with tumor is converted into numerical representation. Then the objective function is executed with minimum threshold value to obtain exact area of tumor. On decreasing threshold value more unique cluster is obtained. Figure.1 shows MRI brain image comparative measures with Fuzzy C-Means and Fuzzy-UCAM.



Figures 3. (a) Normal MRI brain image (b) original MRI with tumor (c) Color map tumor image (d) Tumor image segmented using fuzzy C-Means (e) Segmented image using Fuzzy-UCAM

Figure 3(a) and (b) shows the normal MRI brain image and MRI with tumor where both the images are in gray scale representation. To have a better visualization on tumor the gray scale images are converted into color map image of tumor. Figure 3(d) shows the tumor image which is segmented using fuzzy C-Means where the tumor area is deducted along with the surrounding pixels of the tumor image but it haven't deducted the exact area of the tumor. By increasing the number of cluster *c* in fuzzy C-Means, it increases the number of cluster but there is least possibility of deducting exact area of tumor, since it priorly fixes number of cluster and its seeds randomly. Figure 3(e) segmentation of brain tumor image directly using fuzzy-UCAM, since it uses UCAM clustering algorithm for its initial clustering. Clearly the image shows that the exact area of tumor has been extracted and by adjusting the threshold value, area affected by the tumor can also be extracted. On adjusting the threshold value it increases and decreases the number of cluster but it does not affects the cluster uniqueness, since fuzzy-UCAM purely works on affinity measure. The quality of the segmentation results can be evaluated in terms of segmentation accuracy, which is calculated as follows.

$$A_s = (N_c / T_p) \times 100 \tag{7}$$

where *N_c* is the number of correctly segmented pixels and *T_p* is the total number of pixels in the given image.

D. Efficiency(CPU Time)

In this work, a comparative analysis is performed on the techniques based on the performance measures and based on the deduction exact area of tumor and convergence rate. Consideration one is clearly stated and experimented where as convergence rate is the time period required for the system to reach the stabilized condition. Significant improvement is achieved in Fuzzy-UCAM over Fuzzy C-Means in terms of convergence rate which is clearly visualized in the below notated graph. Better results may achieve if more than two bits are changed in the bit mask.

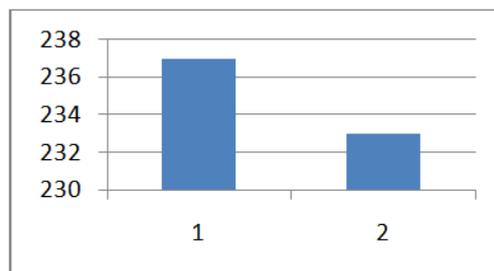


Figure 4. Convergence time of Fuzzy C-Means(1) and Fuzzy-UCAM(2)

E. Comparative Analysis

The comparative study of K-Means, FCM, UCAM and Fuzzy-UCAM clustering are shown in the following table 1.

TABLE I. Comparative study on K-Means, Fuzzy C-Means, UCAM and Fuzzy-UCAM Clustering algorithm

	Initial cluster	Centriod	Threshold value	Cluster Result	Cluster Error
K-Means	K	Initial seeds	-	Depend on initial seeds	Yes, if wrong seeds
Fuzzy C-Means	C	Initial seeds	-	Depend on initial seeds	Yes, if wrong seeds
UCAM	-	-	T	Depend on threshold value	-
Fuzzy-UCAM	-	-	T	Depend on threshold value	-

UCAM and Fuzzy-UCAM algorithm produce unique clustering only on the bases of affinity measure; hence there is no possibility of error in clustering. One major advantage of UCAM and Fuzzy-UCAM algorithm is that both rough clustering and accurate unique clustering is possible by adjusting the threshold value. But in K-Means and FCM clustering there is chance of getting error if the initial seeds are not identified properly.

V. CONCLUSION

In this research paper, we used new UCAM and Fuzzy-UCAM algorithm for segmenting the Magnetic Resonance Image (MRI) with brain tumor and which successively segments tumor through affinity measure without using any prior fixation on number of cluster and random selection of seeds as in K-Means and in fuzzy C-Means. UCAM and Fuzzy-UCAM fixes threshold value to obtain a unique clustering. The proposed methods improve the scalability and reduce the clustering error. This approach ensures that the total mechanism of clustering is in time without loss in correctness of clusters. The visualization and detective valuations of the results of the segmentation show the success of the approaches. The modified UCAM algorithm yields superior convergence rate. The tumor identification and the investigation are carried out for the potential use of MRI data for improving the tumor shape and 2D visualization of the surgical planning.

REFERENCES

- [1] Anton Bardera, Jaume Rigau, Imma Boada, Miquel Feixas, and Mateu Sbert, "Image Segmentation Using Information Bottleneck Method", Page Number 1601-1612, IEEE Transactions on Image Processing, Vol. 18, No. 7, July 2009.
- [2] S. Guha, R. Rastogi, and K. Shim. CURE: An efficient clustering algorithm for large databases. In Proc. 1998 ACM SIGMOD Int. Conf. Management of Data (SIGMOD'98), pages 73–84, 1998.
- [3] Chen Zhang and Shixiong Xia, "K-Means Clustering Algorithm with Improved Initial center," in Second International Workshop on Knowledge Discovery and Data Mining (WKDD), pp. 7906792, 2009.
- [4] F. Yuan, Z. H. Meng, H. X. Zhang, C. R. Dong, "A New Algorithm to Get the Initial Centroids", proceedings of the 3rd International Conference on Machine Learning and Cybernetics, pp. 26629, August 2004.
- [5] Chaturvedi J. C. A, Green P, "K - Modes clustering," Journals of Classification, (18):35–55, 2001.
- [6] Daxin Jiang, Jian Pei, and Aidong Zhang "An Interactive Approach to mining Gene Expression Data". IEEE Transactions on Knowledge and Data Engineering, vol 17, No.10, pp.13636 1380, October 2005
- [7] Doulaye Demele and Philippe Kastner, "Fuzzy C means method for clustering microarray data", bioinformatics, vol.19, no.8, pp.9736 980, 2003.
- [8] Dongxiao Zhu, Alfred O Hero, Hong Cheng, Ritu Khanna and Anand Swaroop, "Network constrained clustering for gene microarray Data", doi:10.1093 bioinformatics / bti 655, Vol. 21 no. 21, pp. 4014 – 4020, 2005.
- [9] G.K. Gupta "Data mining with case studies".
- [10] H. B. Kekre. T. Sarode. and K. Raut. "Detection of Tumor in MRI Using Vector Quantization segmentation". International Journal of Engineering Science and Technology. Vol. 2(8), 2010, pp. 3753-3757.
- [11] A. Wee, C. Liew, and H. Yan. "Current Methods in the Automatic Tissue Segmentation of 3D Magnetic Resonance Brain Images". Current Medical Imaging Reviews, Vol. 2, No. 1, 2006, pp.1-13
- [12] R.C. Gonzalez and R.E. Woods. "Digital Image Processing". Prentice Hall, New Jersey, 2008
- [13] J.C. Rajapakse. J.N. Giedd. and J.L. Rapoport. "Statistical Approach to Segmentation of Single Channel Cerebral MR Images". IEEE Trans. on Medical Imaging. Vol.16. No.2. April 1997
- [14] M.W. Vannier. C.M. Speidel. and D.L. Rickman. "Magnetic resonance imaging multi spectral tissue classification". Journal of NIPS. vol.3. Aug. 1991
- [15] M. Ozkan. B.M. Dawant. and R.J. Maciunas. "Neural Network Based Segmentation of Multi Modal Medical Images. A Comparative and Prospective Study". IEEE Transactions on Medical Imaging. Vol. 12. No.3. Sep. 1993. pp. 534-544.
- [16] M.W. Hansen and W.E. Higgins. "Relaxation Methods for Supervised Image Segmentation". IEEE Trans. on Pattern Analysis and Machine Intelligence. Vol. 19. No. 9. Sep. 1997
- [17] M.N. Ahmed and A.A. Farag. "Two stages Neural Network for Medical Volume mentation". Accepted for Publication in the Journal of Pattern Recognition Letters. 1998
- [18] L.O. Hall. A.M. Bensaid. L.P. Clarke. R.P. Velthuizen. M.S. Silbger. and J.C. Bezdek. "A Comparison of Neural Network and Fuzzy Clustering Techniques in Segmenting Magnetic Resonance Images of the Brain". IEEE Transactions on Neural Networks. Vol. 3. No. 5. Sep.1992. pp. 672-681.

- [19] M.G. Masoole and A. S. Moosavi. "An Improved Fuzzy Algorithm for Image Segmentation". World Academy of Science, Engineering and Technology Vol. 38. 2008. pp. 400-404.



A. Pethalakshmi was born in 1965 at Karaikudi, Tamil Nadu (TN), India. She received her Master of Computer Science from Alagappa University, Tamilnadu, India in 1988 and the Master of Philosophy in Computer Science from Mother Teresa Women's University, Kodaikanal, Tamilnadu, India in 2000. She obtained her Ph.D degree from the Mother Teresa Women's University, Kodaikanal, Tamilnadu, India in 2008. Currently she is working as Associate Professor and Head, Department of Computer Science, M.V.M Government Arts College(W), Dindigul, Tamilnadu, India. Her area of interests includes Data Mining, Rough Set, Fuzzy Set, Grid Computing and Neural Network.



A. Banumathi was born in Karur, Tamil Nadu (TN), India, in 1978. She received the Bachelor of Computer Science (B.Sc.) degree in 1998, Master of Computer Applications (M.C.A.) degree in 2001 and Master of Philosophy in Computer Science in 2005 from Bharathidasan University, Trichy, Tamilnadu, India. She is currently pursuing her Ph.D. degree in Computer Science from Manonmaniam Sundaranar University. Her research interests include data mining, Fuzzy Set and Rough Set.