



# Smallest Univalued Segment Assimilating Nucleus approach to Brain MRI Image Segmentation using Fuzzy C-Means and Fuzzy K-Means Algorithms

AJALA Funmilola Alaba<sup>1</sup>, AKANDE Noah Oluwatobi<sup>2\*</sup>, ADEYEMO Isiaka Akinkunmi<sup>3</sup>,  
Ogundokun Roseline Oluwaseun<sup>4</sup>

<sup>1</sup>Computer Science and Engineering Department, Ladoke Akintola University of Technology, Ogbomosho, Oyo State.

<sup>2\*,4</sup> Computer Science Department, Landmark University, Omuaran, Kwara State.

<sup>3</sup>Computer Science Department, University of Nigeria, Nsukka, Enugu State.

## ABSTRACT

Image segmentation still remains an important task in image processing and analysis. Sequel to any segmentation process, preprocessing activities carried out on the images have a great effect on the accuracy of the segmentation task. This paper therefore laid emphasis on the preprocessing stage of brain Magnetic Resonance Imaging (MRI) images using Smallest Univalued Segment Assimilating Nucleus (SUSAN) and bias field correction algorithms. Subsequently, brain tissue extraction tool was employed in extracting non-brain tissues from the brain image. Afterwards, Fuzzy K-Means (FKM) and Fuzzy C-Means (FCM) segmentation algorithms were employed for segmenting brain MRI images acquired from four different MRI databases into their White Matter (WM), Gray Matter (GM) and Cerebrospinal Fluid (CSF) constituents. Evaluation metrics such as cluster validity functions using partition coefficients and partition entropy; area error metrics such as false positive, true positive, true negative and false negative (FN); similarity index, sensitivity and specificity were used to evaluate the performance of both techniques. A comparative analysis of the experimental results revealed that in most instances, FKM segmentation technique is preferable to FCM segmentation technique for brain MRI segmentation task.

## Indexing terms/Keywords

Brain magnetic resonance imaging, cluster validity functions, fuzzy K-means, fuzzy C-means, segmentation.

## Academic Discipline And Sub-Disciplines

Pattern recognition, Medical Image Processing, Soft Computing

## SUBJECT CLASSIFICATION

Computer Science, Medical Science

## TYPE (METHOD/APPROACH)

Experimental, Computational Analysis

## INTRODUCTION

For accurate information elicitation from medical images, image segmentation is a crucial task that cannot be undermined. It is an essential routine because the eventual outcome of the analysis will determine the success of the pattern image segmentation stage. This involves measuring and visualizing the brain's anatomical structures, analyzing brain changes, delineating pathological regions, surgical planning and image-guided interventions [1]. Generally, image segmentation entails partitioning an image of interest into its constituent regions that are non-overlapping whilst possessing similar attributes that are semantically meaningful. With focus on brain Magnetic Resonance Image (MRI) and iris images, segmentation of MRI image entails classifying the image into its constituent components: the white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF) while segmenting IRIS image entails classifying the Iris image into iris setosa, iris versicolour, and iris virginica. An efficient image segmentation technique must be able to uniquely and correctly classify these dataset into their constituent components to an acceptable degree.

Generally, image segmentation techniques could be classified into layer and block-based segmentation methods [2]. While layered approach defines image shape masks and explain their appearance and depth, the block based approach focuses on extracting features such as colour, textual and spatial information. Block based approach also provides information about the pixels, regions, edges or boundaries of extracted features. This research employs fuzzy variants of C and K means algorithms mostly denoted as Fuzzy C-Means (FCM) and Fuzzy K-Means (FKM) algorithms; a block based approach in segmenting brain MRI and IRIS images. The fuzzy nature of fuzzy algorithms is established in the fact that when an image is broken down into segments, each segment can belong to more than one cluster, though different segments have varying membership function which indicates their strength of association between the element and other clusters. Fuzzy segmentation technique involves assigning membership levels to the segments and then using these levels to assign data elements to one or more clusters. The advantages of fuzzy segmentation approach includes yielding regions more homogeneous than other methods; reducing the spurious blobs; removing noisy spots; reduced sensitivity to noise compared to other techniques [3]. However, fuzzy approach requires prior knowledge about the number of clusters present or possible in a particular dataset; this may be difficult to obtain a priori in a new dataset. In this research work,

emphasis will be laid on Fuzzy C-Means (FCM) and Fuzzy K-Means (FKM) clustering algorithms for brain MRI segmentation.

## RELATED WORKS

FCM has been widely used for image segmentation in several fields such as feature extraction, pattern and image recognition, fuzzy identification etc. it is widely accepted because it preserves more information from the original image than other segmentation methods [4]. FCM segmentation technique was adopted in [5] for the identification and segmentation of lung nodules in Computed Tomography (CT) images. Their results showed that FCM was able to effectively separate the parenchyma from the mediastinum and the thoracic wall from the lung nodules. A study to determine the most suitable cluster validity index needed to achieve optimal clustering was carried out in [3]. The study employed the topmost 18 validity index in turn to synthetic images corrupted with noise of varying levels, and simulated volumetric MRI datasets. The results obtained showed that the various indexes have different outcomes for various noise levels, while some indexes can help in determining the accurate number of clusters existing in a dataset. As no algorithm is entirely flawless, several improvements of FCM have been proposed by researchers, these have led to the emergence of various versions of FCM. Some of these variations reported in [6] include Bias Corrected Fuzzy c-Means (BCFCM), Possibilistic Fuzzy C-Means (PFCM), Spatial Fuzzy C-Means (SFCM), Fuzzy Local Information C- Means (FLICM), Multi-dimensional Fuzzy C-Means (MDFCM), FCM-S1, FCM-S2, Tsallis Entropy Based Fuzzy C-Means (TEFCM), Robust Fuzzy C-means Based Kernel Function (RFCMK), Weighted Image Patch Based FCM (WIPFCM) and Kernel Weighted Fuzzy Local Information C- Means (KWFLICM). A comprehensive performance analysis among these variations was carried and their performance was validated on real MRI images as well as synthetic images. Partition coefficient ( $V_{pc}$ ), partition entropy ( $V_{pe}$ ), time complexity and segmentation accuracy were used as the performance evaluation metrics. High value of  $V_{pc}$  implies better performance while low value of  $V_{pe}$  implies better clustering. SFCM, TEFCM and MDFCM yielded a high  $V_{pc}$  and low  $V_{pe}$  values which denotes better performance and clustering. When considering feature structure, RFCMK yielded an appreciable performance while WIPFCM, MDFCM and KWFLICM also produce a better performance. Fuzzy c means was modified for segmenting T1-T2 weighted brain MRI images in [4], the modified algorithm called Modified Robust Fuzzy c-Means with weight Bias Estimation (MRFCM-wBE) was employed to solve the intensity inhomogeneity and noise affecting the segmentation process of computer assisted segmentation algorithms across differentiating borders between tissues of medical images. Fuzzy C-means was modified by reducing the number of iterations using dist-max initialization algorithm after which the algorithm was executed iteratively. Running time for obtaining clusters, number of iterations for completion of clustering the datasets, and silhouette width for clustering accuracy were the yardstick used to measure the performance of MRFCM-wBE. Checkerboard and lung cancer datasets were used as test data. The performance of the algorithm was compared with other segmentation algorithms such as Gaussian Kernel based Fuzzy c-Means algorithm (GKFCM), BCFCM, Improved Fuzzy Segmentation (IFS) algorithm and SFCM. The algorithm showed an appreciable improvement in running time over the remaining algorithms. In [7], Micro calcification clusters in digitized mammograms were detected using FCM and Possibilistic FCM segmentation techniques. The approach adopted involved detecting and separating micro calcifications which are small deposits of calcium in breast tissue. The results obtained showed that the algorithms adopted could detect and separate micro calcifications for normal tissues though PFCM performed better since its obtained results depend solely on the threshold value being the most appropriate. FCM has also been used in [8] for the detection of leukemia from blood images. This employs morphological contour segmentation approach to identify the edges of nuclei in white blood cells and further separates them from the microscopic blood image. The approach adopted yielded a better result in terms of accuracy and time consumption when compare to normal hematologist's visual classification. In addition to leukemia detection, FCM segmentation technique could be used in jaw tumor detection as carried out in [9] using panoramic X-ray images as test data. The approach employed involved grouping and clustering jaw's X-ray image pixels into background region and lesion region. Comparing the results obtained with other segmentation techniques, an appreciable result was obtained in terms of specificity, Similarity Index, Area error metrics and Sensitivity value.

Fuzzy K-Means (FKM) – a variant of k-means algorithm has also been widely used for image segmentation. In FKM segmentation technique, the data of interest is partitioned into k-groups of disjoint cluster. To begin with, the k-centroid point is calculated after which the cluster with the nearest distance to the centroid point is determined. A research to determine if FKM actually divides a data into k groups and if the k-clusters chosen is accurate was carried out in [10]. Their work also looks into the color space at which KCM segments data better. Experiments were carried out on  $L^*a^*b$  colour space and RGB colour space images using 1, 2 and k-cluster color images ( $k > 3$ ). Silhouette Width which is a function of cluster's tightness and separation likewise Ground truth images were used to evaluate the adopted approach. Results showed that KCM considerably segment images much better in  $L^*a^*b^*$  colour space when compared to RGB colour space.

The initial clusters centre in FKM algorithms are generated randomly; this leads to non-repeatable clustering results that may be hard to comprehend. As a result of this, a kernel density based clustering technique where the data is mapped to a high-dimensional space for classification was proposed in [11]. This involved selecting an initial seed value with maximum variance from the data matrix, then employing Gaussian kernel for the estimation of the initial density. Afterwards, subsequent seed values are selected that has densities equivalent to the initial selected points. In this way, k points with similar density with respect to each other are obtained. The proposed technique was able to avoid the dead and trapped centre limitation of clustering based segmentation approach.

The performance of three fuzzy algorithms viz: FCM, Gustafson–Kessel (GK) and FKM algorithms were analyzed based on their clustering output criteria in [12]. Liver disorders and wine datasets were used to test the performance of these algorithms. With FKM and FCM using Euclidian distance measure and GK using fuzzy covariance matrix in their distance

metrics, the obtained experimental results showed that GK produced a similar result to FCM but FKM outperforms both FCM and GK algorithms. Thus, the efficiency of FKM was reported to be better than FCM and GK algorithms. Few comparative work between FKM and FCM have been carried out so far, hence, this research work attempts to further test FKM and FCM algorithm on brain MRI images with a view to establish which algorithms perform better.

## Materials and Methods

The proposed method consists of pre-processing and the actual segmentation stage. The result of the segmentation stage to a large extent is dependent on the success of the preprocessing stage; hence, efforts were put in place to ensure that the pre-processing stage was meticulously carried out.

### Pre-processing stage

The preprocessing stage involves three steps which are: noise reduction, bias field correction and brain extraction.

#### Noise Reduction

To eliminate the effect of noise on the image and improve local features using local smoothing, Smallest Univalued Segment Assimilating Nucleus (SUSAN) algorithm illustrated in [13] was used. SUSAN algorithm perfectly works for carrying out edge detection, corner detection and structure-preserving image noise reduction. SUSAN algorithm achieves this by placing a predetermined window (mask) at each point of an image and then calculates the brightness threshold of each pixel within the mask at each point. The calculated brightness was compared with the center point or nucleus of the pixel using equation (1):

$$c(\vec{r}, \vec{r}_0) = \begin{cases} 1 & \text{if } |I(\vec{r}) - I(\vec{r}_0)| \leq t \\ 0 & \text{if } |I(\vec{r}) - I(\vec{r}_0)| > t \end{cases} \quad (1)$$

Where  $\vec{r}_0$  is the position of the nucleus in the image and  $\vec{r}$  is the position at any other point within the mask.  $I(\vec{r})$  is the brightness of any pixel with "t" as the brightness difference threshold and "c" is the output of the comparison. This comparison was done for each pixel within the mask using equation (2):

$$n(\vec{r}_0) = \sum_{\vec{r}} c(\vec{r}, \vec{r}_0) \quad (2)$$

Where n is the number of pixels in the USAN. The comparison gives the USAN's area which will be ultimately minimized. The brightness difference threshold "t" helps in knowing the minimum contrast of features which will be detected likewise the maximum amount of noise which will be ignored. The next step of SUSAN algorithm is to carry out a comparison between "n" and a fixed threshold "g". There would have been no need for the geometric threshold if noise were to be absent. Nonetheless, to ensure noise is totally eradicated or reduced to the nearest minimum, "g" is set to  $3n_{\max}/4$  where

$n_{\max}$  is the maximum value of "n". The initial edge response  $R(\vec{r}_0)$  is calculated using equation (3):

$$R(\vec{r}_0) = \begin{cases} g - n(\vec{r}_0) & \text{if } n(\vec{r}_0) < g \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

Depending on the type of edge point being expected (typical straight edge or standard edge points), the direction of the edges associated to an image point must be calculated. If it were to be a typical straight edge, equation (3) can be used to compute the required edge else the direction of the edge must be calculated by first computing the center of gravity followed by finding the longest axis of symmetry. The center of gravity could be computed using equation (4) while the edge direction is projected by finding the sum of equations (5, 6 and 7):

$$\vec{r}(\vec{r}_0) = \frac{\sum_{\vec{r}} \vec{r} c(\vec{r}, \vec{r}_0)}{\sum_{\vec{r}} c(\vec{r}, \vec{r}_0)} \quad (4)$$

$$\overline{(x - x_0)^2}(\vec{r}_0) = \sum_{\vec{r}} (x - x_0)^2 c(\vec{r}, \vec{r}_0) \quad (5)$$

$$\overline{(y - y_0)^2}(\vec{r}_0) = \sum_{\vec{r}} (y - y_0)^2 c(\vec{r}, \vec{r}_0) \quad (6)$$

$$\overline{(x - x_0)(y - y_0)}(\vec{r}_0) = \sum_{\vec{r}} (x - x_0)(y - y_0) c(\vec{r}, \vec{r}_0) \quad (7)$$

If the computed USAN area is smaller than the mask diameter then the edge belongs to the typical edge category however, if the USAN area is larger then, it belongs to the standard edge category [1].

#### Bias Field Correction

Bias field arises as a result of spatial inhomogeneity of the magnetic field, deviations in the sensitivity of the reception coil, and contact between the magnetic field and the human body [22]. Scanning MRI at 0.5T could make the effect of bias field minimal, hence, its effect could be neglected but when MRI are scanned with MR scanners with magnetic field above

1.5T, the effect of the bias field is always strong and could affect MRI analysis [4], at this instance the effect needs to be corrected. To correct the effect of this signal, the modified fuzzy-c means approach reported in [21] was adopted. This entails the following steps:

- (a). Choosing an initial cluster prototypes  $\{v_i\}_{i=1}^c$  and setting the bias field  $\{b_k\}_{k=1}^n$  to a very small positive value
- (b). Updating partition matrix  $\|u_{ij} = U\|$  where  $U \in U$  using equation (8):

$$u_{ij} = \frac{1}{\sum_{j=1}^c \left( \frac{D_{ik} + \frac{\alpha}{N_R} \gamma_i}{D_{ik} + \frac{\alpha}{N_R} \gamma_i} \right)^{1/m-1}} \quad (8)$$

Where  $D_{ik} = \|y_k - B_k - v_i\|^2$  and the distance expression  $\gamma_i = \left( \sum_{y_k \in N_k} \|y_k - b_k - v_i\|^2 \right)$ .  $y_k$  is the observed intensity at the  $k_{th}$  pixel,  $b_k$  is the bias field at the  $k_{th}$  pixel,  $v_i$  is the  $i_{th}$  cluster prototype,  $N_k$  is the set of neighbours that exist in a window around the true intensity,  $N_R$  is the cardinality of this set of neighbours and the effect of these neighbours is controlled by the parameter  $\alpha$ .

- (c). Obtaining new prototypes of the clusters  $v_i^*$  in the form of weighted averages of the patterns using:

$$v_i^* = \frac{\sum_{k=1}^N u_{ik}^m (y_k - \beta_k) + \frac{\alpha}{N_R} \sum_{y_k \in N_k} (y_k - \beta_k)}{(1+\alpha) \sum_{k=1}^N u_{ik}^m} \quad (9)$$

- (d). Compute the bias term using:

$$\beta_k^* = y_k - \frac{\sum_{i=1}^c u_{ik}^m v_i}{\sum_{i=1}^c u_{ik}^m} \quad (10)$$

- (e). Steps (b) – (d) will be repeated till termination criteria in (11) is satisfied

$$\|V_{new} - V_{old}\| < \epsilon \quad (11)$$

Where  $V$  is a vector of cluster centers and  $\epsilon$  is a small number that can be taken during the initialization process.

## Removal of Non-brain Tissue

Since emphasis is on brain tissue segmentation, removal of non-brain and extra-cerebral tissues like skull, neck, fat, bones, skin, eyeballs e. t. c. is necessary. These tissues can distort the segmentation process and affect the accuracy of the segmentation result. There are several brain tissue extraction algorithms such as: statistical parametric mapping, brain surface extractor, minneapolis consensus strip, threshold morphologic brain extraction and Brain Extraction Tool (BET). This research work employs BET; a widely known publicly available brain tissue extraction tool in extracting non-brain tissues from the brain image. The “generate binary brain mask image” option of BET was used to effect the brain tissue extraction.

## Segmentation Stage

The segmentation task was carried out using FCM and FKM clustering algorithms.

## FUZZY C-MEANS Technique

With FCM, a pattern or data sample can belong to more than one cluster while the membership value assigned to each pattern is used to indicate the similarity. In this research, FCM clustering begins by supplying a set of data and a number of clusters into Matlab environment using the Matlab function fcm. The function in turn returns the best cluster centers and membership value for each pattern. FCM clustering is an iterative process, hence, the cluster centers was repeatedly moved to the right location within the dataset by minimizing the objective function. The iteration terminates when the objective function enhancement between two consecutive iterations is less than the minimum amount of improvement specified [16]. FCM algorithm as illustrated in [3, 4, 17] expects a dataset of interest “x” with n data samples such that  $x = \{x_1, x_2, \dots, x_n\}$  to be partitioned into clusters “c”. FCM clustering task begins by minimizing the objective function  $J$  which is defined as:

$$J_m = \sum_{i=1}^n \sum_{j=1}^c u_{ij}^m \|x_i - c_j\|^2 \quad (12)$$

Given  $1 \leq m \leq \infty$ ,  $m$  is the fuzziness constant (any positive number greater than 1) which decides the fuzziness of the resulting partitions.  $u_{ij}$  is the membership value of dataset  $x_i$  for each class "l" and pixel "j",  $c_j$  is the  $j^{\text{th}}$  cluster while  $n$  is the number of clusters. Since FCM employs an iterative process, this is achieved by with updating the membership value  $u_{ij}$  and cluster center  $c_j$  repeatedly using equations (13) and (14): for each class is assigned to each pixel ,

$$u_{ij} = \frac{1}{\sum_{n=1}^c \left( \left\| \frac{x_i - c_j}{x_i - c_n} \right\| \right)^{\frac{2}{m-1}}} \quad (13)$$

$$c_j = \frac{\sum_{i=1}^n u_{ij}^m \cdot x_i}{\sum_{i=1}^n u_{ij}^m} \quad (14)$$

Where "·" is a norm metric. The membership value  $u_{ij}$  must satisfy the constraint:

$$0 \leq u_{ij} \leq 1 \quad (15)$$

$$0 < \sum_{k=1}^n u_{ij} < n \quad (16)$$

$$\sum_{i=1}^c u_{ij} = 1 \quad (17)$$

Where  $1 \leq i \leq c, 1 \leq j \leq n$  for equation (15),  $1 \leq i \leq c$  for equation (16) and  $1 \leq j \leq c$  for equation (17). The iteration phenomenon continues until equation (18) is satisfied:

$$\max_{ij} \left\{ \left| u_{ij}^{(k+1)} - u_{ij}^{(k)} \right| \right\} < \varepsilon \quad (18)$$

## FUZZY K-MEANS Technique

Integrating fuzzy logic into K-Means clustering algorithm yields FKM [16]. FKM as illustrated in [13, 15] implies that the data set of interest well be partitioned into  $k$ -clusters  $S_i = (1, 2, 3, \dots, k)$  such that each cluster  $S_i$  is compact and far from other clusters. This can be achieved by minimizing the distance between the cluster centers  $c_j$  and the patterns that belong to that cluster. To establish the fuzziness of the pattern and the cluster center, the degree of the belongingness was represented by a characteristic function  $u_{ij} = 0$ . This function  $u_{ij}$  must fulfill the constraint:

$$\sum_{j=1}^k u_{i,j} = 1 \quad (19)$$

The objective function  $J$  must be minimized with respect to  $c_j$  and  $u_{ij}$  using equation (20), such that

$$J = \sum_{j=1}^k \sum_{i=1}^N u_{i,j} d_{i,j} \quad (20)$$

Where the number of patterns is denoted by  $N$ , the measured Euclidean distance between  $S_i$  and  $c_j$  is denoted by  $d_{ij}$  while the number of clusters is denoted by  $k$ . Typically, the patterns are not expected to overlap but if they do, then each pattern belongs to more than one cluster i.e.  $u_{ij} = [0, 1]$ . In this instance,  $u_{ij}$  should be referred to as a membership function rather than a characteristic function. As a result of this, equation (20) was modified as

$$J = \sum_{j=1}^k \sum_{i=1}^N u_{i,j}^m d_{i,j} \quad (21)$$

Where  $m$  remains the fuzzifier parameter which controls the degree of fuzziness. Minimizing the objective function  $J$  may produce an insignificant solution, hence, constraints illustrated in equations (22) and (23) must be adhered to

$$\sum_{i=1}^N u_{i,j}^m > 0 \text{ where } i, j=1 \dots k \quad (22)$$

$$\sum_{j=1}^k u_{i,j}^m = 1 \text{ where } i, j= 1 \dots N \quad (23)$$

Constraint in equation 22 ensures that there is no empty cluster while the constraint in the equation 23 enforces a condition that each pattern must share its membership with all other clusters such that the sum of membership equals one [15]. The memberships and the cluster centers need to be updated until the changes in the membership values of all the patterns become insignificant or the required number of iterations has been reached. The update was achieved by using equation (24):

$$u_{i,j} = \left( (d_{i,j})^{1/m-1} \sum_{l=1}^k \left( \frac{1}{d_{il}} \right)^{1/m-1} \right)^{-1} \quad (24)$$

If  $d_{ij} < \eta$ ,  $u_{ij} = 1$  where  $\eta$ ; a very small positive number is taken to be 0.01. To obtain a new set of cluster centers, equation (25) was employed:

$$c_j(p) = \frac{\sum_{i=1}^n u_{ij}^m x_i}{\sum_{i=1}^n u_{ij}^m} \quad (25)$$

The iteration terminates when  $\|c_j(p) - c_j(p+1)\| < \epsilon$  where  $j = 1 \dots k$ , else the iteration continues with  $p = p + 1$ .  $\epsilon$  remains a very small positive number (0.01)

All algorithms were implemented in MATLAB 7.12 (R2011a) programming environment. It was run on an Intel(R) 2.50GHz Core(TM) i5-3210M CPU, 4.00 GB RAM personal computer under windows 8 operating system.

## EXPERIMENTAL RESULTS

### Dataset used

The segmentation approach was validated using four different brain MRI datasets. Real brain MRI datasets were downloaded from [18]. The dataset was acquired using a physical phantom on a 3T MRI scanner with a turbo spin-echo sequence. The data was acquired with a 220 mm x 292 mm field of view on a 256x340 Cartesian sampling grid and was collected with an array of receiver coils, leading to 16 channels of information. The data was provided in Matlab format and were referred to as "D1". Synthetic brain MRI dataset was also downloaded from [23]. Each image in the downloaded dataset (referred to as D2) has a slice thickness of 1mm, resolution of 1mmX1mmX1mm, 20% intensity non-uniformity (RF), noise levels ranging from 0% to 9%, and 181X217X181 matrix voxels. All these parameters are provided before the final dataset is downloaded. Manually labeled brain MRI images [D3] were also downloaded from the Internet Brain Segmentation Repository (IBSR) [24]. The labelled images were not created by an algorithm but by neuroanatomical experts. The images in the datasets have different levels of difficulty which can be used to evaluate the segmentation task in different conditions. Finally, T1 and T2 brain MRI dataset from [25] with 1x1x1 mm<sup>3</sup> resolution and a voxel size of 2x2x2 mm<sup>3</sup> were also used.

### Evaluation Metrics

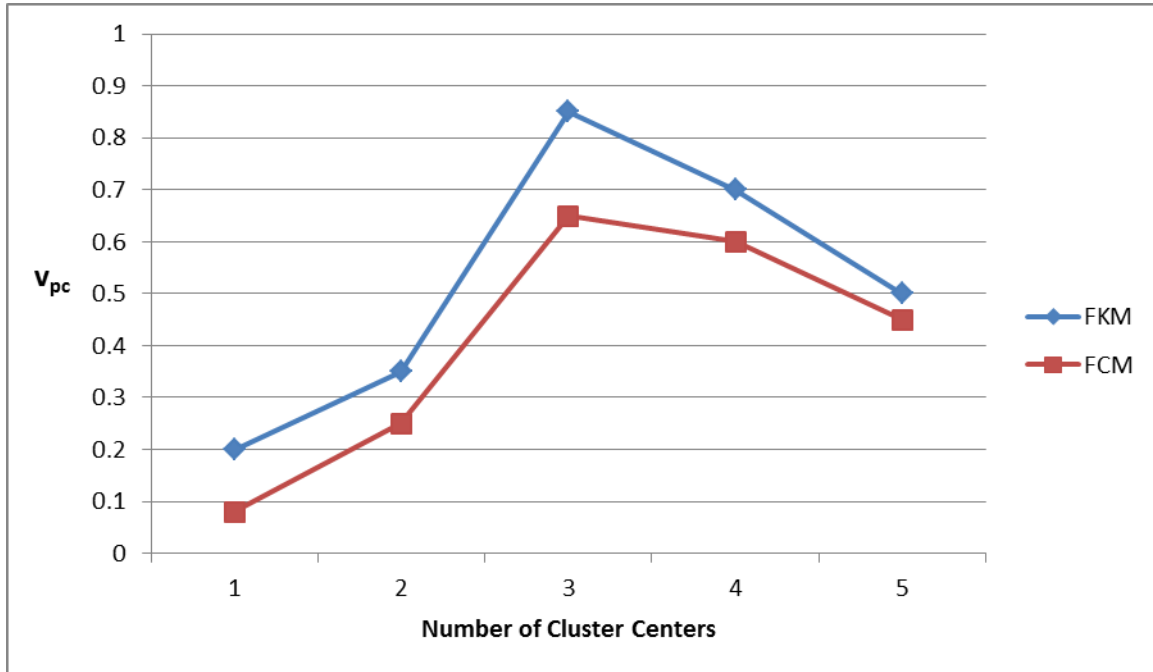
The performance of both FCM and FKM algorithms were measured using various performance evaluation metrics such as cluster validity functions, Area error metrics, similarity index, sensitivity and specificity

## Cluster Validity Function

Cluster validity function uses fuzzy partition to determine the efficiency of the segmentation task. Fuzzy partition is a function of the partition coefficient  $v_{pc}$  and partition entropy  $v_{pe}$ .

**Partition Coefficients:** Partition coefficients ranges from 0 to 1 with a high value indicating a better performance. It is computed using equation (26)

$$v_{pc} = \frac{\sum_{j=1}^N \sum_{i=1}^c u_{ij}^2}{N} \quad (26)$$



**Fig 1: Fuzzy Partition Coefficient using D1**

As shown in Figure 1, the results showed that both FKM and FCM could correctly classify the phantom dataset into three clusters but at different partition coefficient. FKM having the highest partition coefficient 0.85 at the third cluster center shows that FKM segments the dataset better at a higher partition coefficient when compared to FCM at 0.65. Partition coefficients of the remaining three datasets are highlighted in table 1. The results show that both techniques have the peak partition coefficients at the 3<sup>rd</sup> cluster points across all the datasets (D2, D3 and D4) used. This showed that all the algorithms could at least identify the three clusters correctly.

**Table 1. Partition coefficients across three datasets D2, D3, D4**

Number of Clusters	D2		D3		D4	
	V <sub>pc</sub>		V <sub>pc</sub>		V <sub>pc</sub>	
	FKM	FCM	FKM	FCM	FKM	FCM
1	0.25	0.20	0.28	0.26	0.24	0.22
2	0.37	0.36	0.38	0.36	0.36	0.32
3	0.82	0.80	0.81	0.79	0.80	0.76
4	0.76	0.72	0.74	0.73	0.76	0.75
5	0.48	0.45	0.47	0.44	0.47	0.46

## Partition Entropy

A high partition coefficient is expected for an accurate clustering while low partition entropy is expected. It is computed using equation (27) while the results across the four datasets D1, D2, D3 and D4) are shown in Figures 2-5.

$$v_{pe} = \frac{-\sum_{j=1}^N \sum_{i=1}^C u_{ij}^2}{N} \quad (27)$$



Figure 2: Effect of Partition Entropy on Dataset "D1"

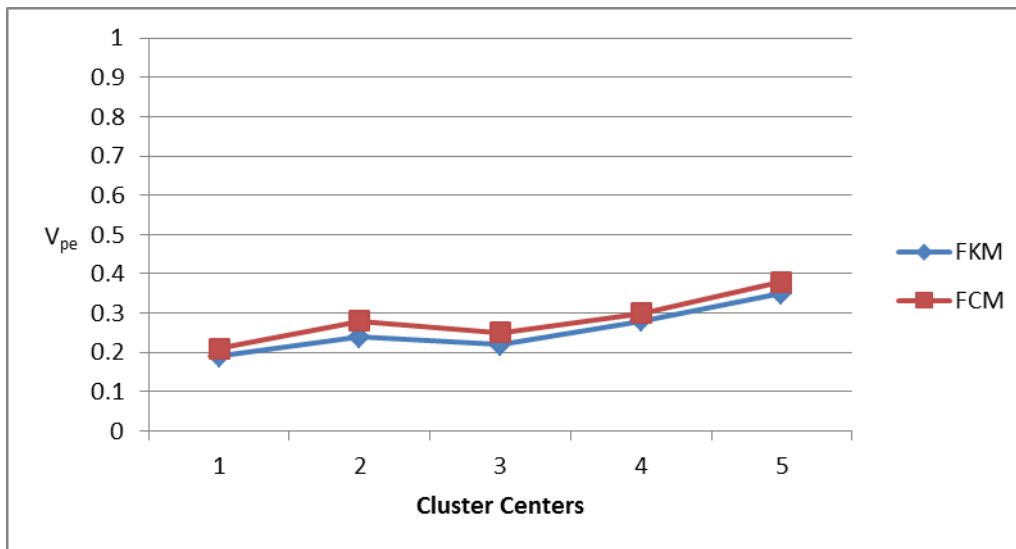


Figure 3: Effect of Partition Entropy on Dataset "D2"





Figure 4: Effect of Partition Entropy on Dataset "D3"



Figure 5: Effect of Partition Entropy on Dataset "D4"

### Area Error Metrics

Using manually labeled brain MRI images [D3] as the reference image, the area ratio of the False Positive (FP), True Positive (TP), True Negative (TN) and False Negative (FN) as used in [9] was employed to determine how correctly FKM and FCM could classify dataset D3 into its constituent WM, GM and CSF component. The area metrics are given as:

$$TP = \frac{|A_m \cap A_a|}{A_m} \quad (28)$$

$$FP = \frac{|A_m \cup A_a - A_m|}{A_m} \quad (29)$$

$$FN = 1 - TP = \frac{|A_m \cup A_a - A_a|}{A_m} \quad (30)$$

The manually measured cluster area is represented by  $A_m$  while those computed using FCM and FKM are represented by  $A_a$ . The results obtained as illustrated in Table 2 showed that FKM performed preferably better when compared with FCM on dataset "D3". FKM having a high TP at 94% when compared to that of FCM at 90% showed that FKM could identify more area for each clusters than FCM. Also, the FP of FKM being lower than FCM showed that FKM was able to

recognize clusters accurately than FCM. The P-value of each metric is less than 0.005 and this statistically showed that the result is highly significant.

**Table 2. FKM and FCM Area Error Metrics**

Algorithms	FP (%)	TN (%)	TP (%)	FN (%)
FKM	6.5±4.2	93.5±4.2	82±3.6	18±3.6
FCM	10.0±7.1	90±7.1	77±6.2	23±6.2
P-value	<0.005	<0.005	<0.005	<0.005

### Similarity Index, Sensitivity and Specificity

To further determine the effectiveness of the algorithms, Similarity Index (SI), sensitivity and specificity of each technique were computed using the following equations:

$$SI = \frac{A_m \cap A_d}{A_m \cup A_d} \times 100 \quad (31)$$

$$Specificity = \frac{TN}{TN+FP} \quad (32)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (33)$$

SI could also be used to know if there is an overlap among the clusters or not. Specificity can be used to determine if an algorithm could classify an image correctly into the expected clusters (WM, GM, CSF in this sense) while sensitivity calculates the number of images that were correctly classified. The mean of the computed metrics on D3 is as described in Table 3. The result showed that FKM segmentation technique is more accurate and reliable in WM, GM and CSF segmentation when compared to FCM, though the results of FCM segmentation is not totally bad.

**Table 3. SI, Sensitivity and Specificity Mean values for FKM and FCM**

Algorithm	Evaluation Metrics (%)	GM	WM	CSF
FKM	SI	99.20	99.26	99.31
	Sensitivity	99.00	98.95	98.75
	Specificity	99.10	99.21	99.18
FCM	SI	97.85	97.62	97.71
	Sensitivity	98.55	98.63	98.76
	Specificity	98.68	98.72	98.82

### CONCLUSION

A comparative analysis aimed at asserting a superior segmentation technique between FKM and FCM algorithm has been carried out in this paper. Using partition coefficients as a metric, both techniques have the peak partition coefficients at the 3<sup>rd</sup> cluster points across all the datasets. This showed that the techniques could correctly classify the datasets into three clusters, though, FKM records a higher value when compared to FCM across all the datasets used which showed that FKM classifies at a faster rate when compared to FCM. Also, when the partition entropy for each technique was computed, FKM attained lower partition entropy when compared to FCM and this further established the superiority of FKM segmentation technique. With the manually labeled brain MRI images as a reference image, the area error metrics (FP, TN, TP, FN) of the techniques were computed to further know the technique that classifies the datasets better. FKM achieved good results which further establish the fact that it's a better technique when compared to FCM. The P-value of each metric being less than 0.005 statistically showed that the computed result is highly significant. Conclusively, similarity index, sensitivity and specificity of each technique using dataset D3 were computed. The mean of the computed metrics revealed that FKM segmentation technique is more accurate and reliable in WM, GM and CSF segmentation when compared to FCM. Based on the results of the various evaluation metrics used, it could be conjectured that FKM is a good segmentation technique and its preferable to FCM segmentation technique.



## References

1. Nida M. Zaitoun, Musbah J. Aqel (2015), "Survey on Image Segmentation Techniques", proceedings of International Conference on Communication, Management and Information Technology (ICCMIT 2015), Procedia Computer Science, Vol. 65, pp. 797 – 806.
2. Ivana Despotović, Bart Goossens, and Wilfried Philips (2015), "MRI Segmentation of the Human Brain: Challenges, Methods, and Applications", Computational and Mathematical Methods in Medicine, Volume 2015, pp 1-23
3. S. Ramathilagam, R. Pandiyarajan, A. Sathya, R. Devi and S.R. Kannan (2011), "Modified fuzzy c-means algorithm for segmentation of T1–T2-weighted brain MRI", Journal of Computational and Applied Mathematics, Vol. 23, No. 5, Pp. 1578–1586
4. Mahipal Singh Choudhry and Rajiv Kapoor (2016), "Performance Analysis of Fuzzy C-Means Clustering Methods for MRI Image Segmentation", Procedia Computer Science, Vol. 89, pp. 749 – 758
5. J. Quintanilla-Dominguez, B. Ojeda-Magaña, M.G. Cortina-Januchs, R. Ruelas, A. Vega-Corona, D. Andina (2011), "Image segmentation by fuzzy and possibilistic clustering algorithms for the identification of microcalcifications, Scientia Iranica, Vol. 18, issue 3, pp. 580–589.
6. Ezhil E. Nithila and S.S. Kumar (2016), "Segmentation of lung nodule in CT data using active contour model and Fuzzy C-mean clustering, Alexandria Engineering Journal
7. E.A. Zanaty (2012), "Determining the number of clusters for kernelized fuzzy C-means algorithms for automatic medical image segmentation", Egyptian Informatics Journal, Vol. 13, pp. 39–58
8. Viswanathan P (2015), "Fuzzy C means Detection of Leukemia based on Morphological Contour Segmentation", Procedia Computer Science, vol. 58, pp. 84 – 90
9. Mutasem K. Alsmadi (2016), "A hybrid Fuzzy C-Means and Neutrosophic for jaw lesions segmentation, Ain Shams Engineering Journal, <http://dx.doi.org/10.1016/j.asej.2016.03.016>
10. S. M. Aqil Burney and Humera Tariq (2014), "K-Means Cluster Analysis for Image Segmentation", International Journal of Computer Applications, Vol.96, No.4, Pp. 0975 – 8887.
11. Ajay Kumar and Shishir Kumar (2016), "Color Image Segmentation via Improved K-Means Algorithm", International Journal of Advanced Computer Science and Applications, Vol. 7, No. 3, Pp. 46-53.
12. B. Simhachalam and G. Ganesan (2016), "Performance comparison of fuzzy and non-fuzzy classification methods", Egyptian Informatics Journal, Vol. 17, pp. 183–188.
13. CHIH-TANG CHANG, JIM Z. C. LAI and MU-DER JENG (), "A Fuzzy K-means Clustering Algorithm Using Cluster Center Displacement", Journal of Information Science and Engineering, Vol. 27, 995-1009.
14. D. Liu and F. Kubala (2004), "Online speaker clustering," in Proceedings of IEEE Conference on Acoustic, Speech, and Signal Processing, Vol. 1, 2004, pp. 333-336.
15. Manish Sarkar and Tze-Yun Leong (2001), "Fuzzy K-means Clustering with Missing Values", In: Proceedings of American Medical Informatics Association, (AMIA), pp. 588–592.
16. Soumi Ghosh and Sanjay Kumar Dubey (2013), "Comparative Analysis of K-Means and Fuzzy C-Means Algorithms", International Journal of Advanced Computer Science and Applications, Vol. 4, No.4, pp. 35-39.
17. Alsmadi M. K. (2015), "MRI brain segmentation using a hybrid artificial bee colony algorithm with fuzzy-C mean algorithm", Journal of Applied Sciences, Vol. 15, pp. 100–109.
18. Real brain MRI dataset downloaded from <http://mr.usc.edu/download/data/> under the [CC Attribution-Noncommercial-Share Alike 3.0 License](#) sponsored by NSF grant CCF-1350563, accessed on 10<sup>th</sup> Nov., 2016.
19. Lichman, M. (2013), "UCI Machine Learning Repository", [<http://archive.ics.uci.edu/ml/>]. Irvine, CA: University of California, School of Information and Computer Science. Accessed on 12<sup>th</sup> Nov., 2016.
20. S.M. Smith and J.M. Brady (1997), "SUSAN - a new approach to low level image processing", International Journal of Computer Vision, vol. 23, No. 1, Pp. 45-78
21. Mohamed N. Ahmed, Sameh M. Yamany, Nevin Mohamed, Aly A. Farag and Thomas Moriarty (2002), "A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data", IEEE Transactions On Medical Imaging, Vol. 21, No. 3, Pp. 193-199.
22. C. M. Collins, W. Liu, W. Schreiber, Q. X. Yang, and M. B. Smith (2005), "Central brightening due to constructive interference with, without, and despite dielectric resonance," Journal of Magnetic Resonance Imaging, vol. 21, no. 2, pp. 192–196
23. Synthetic brain MRI dataset downloaded from <http://brainweb.bic.mni.mcgill.ca/cgi/brainweb1.accessed> on 11<sup>th</sup> August, 2016



24. Manually labeled brain MRI images datasets downloaded from Internet Brain Segmentation Repository (IBSR)-  
<https://www.nitrc.org/projects/ibsr> accessed on 11th August, 2016
25. Bullitt E, Zeng D, Gerig G, Aylward S, Joshi S, Smith JK, Lin W, Ewend MG (2005) Vessel tortuosity and brain tumor malignancy: A blinded study. *Academic Radiology* 12:1232-1240



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

**DOI: 10.24297/ijct.v16i7.6170**