

# Enhanced Image Segmentation: Merging Fuzzy K-Means and Fuzzy C-Means Clustering Algorithms for Medical Applications

Karim Mohammed Aljebory, Thabit Sultan Mohammed\*, Mohammed U. Zainal

Computer Technical Engineering Department, Al-Qalam University College, Kirkuk, Iraq

Received April 28, 2021; Revised May 21, 2021; Accepted June 25, 2021

## Cite This Paper in the following Citation Styles

(a): [1] Karim Mohammed Aljebory, Thabit Sultan Mohammed, Mohammed U. Zainal , "Enhanced Image Segmentation: Merging Fuzzy K-Means and Fuzzy C-Means Clustering Algorithms for Medical Applications," *Computer Science and Information Technology*, Vol. 9, No. 1, pp. 1 - 13, 2021. DOI: 10.13189/csit.2021.090101.

(b): Karim Mohammed Aljebory, Thabit Sultan Mohammed, Mohammed U. Zainal (2021). *Enhanced Image Segmentation: Merging Fuzzy K-Means and Fuzzy C-Means Clustering Algorithms for Medical Applications*. *Computer Science and Information Technology*, 9(1), 1 - 13. DOI: 10.13189/csit.2021.090101.

Copyright©2021 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

**Abstract** Separating an image into regions according to some criterion is called image segmentation. This paper presents an algorithm that combines the fuzzy k-means (FKM) and fuzzy c-means (FCM) clustering strategies. The proposed algorithm combines the FKM and the FCM algorithms mathematical features, which is referred to as (CFKCM). The FKM and FCM clustering algorithms are adopted to compare the performance and hence evaluate the proposed clustering algorithm. Tests are conducted, and performance parameters are calculated for validation. The comparison and assessment analysis are based on metrics related to the image clustering process, such as the Segmentation Accuracy (SA), Clustering Fitness (CF), and cluster Validity function ( $V_{pc}$  and  $V_{pe}$ ). A dataset of MR images is used by this research for the application, test, and evaluation of the image clustering. The results for clustering backbone MRI images show that CFKCM algorithm is featured by being more effective, and comparatively independent of the noise, where it can process the "clean" images and noisy images without knowing the type of the noise, which is the most difficult task in image segmentation.

**Keywords** Image Processing, Clustering, Fuzzy C-Means, Medical Image Segmentation

## 1. Introduction

The field of image processing involves a number of major topics that are considered as hot research areas. Image segmentation is a clear example of these research topics, and few motivations are lying behind such research interests. Every stage of image segmentation is required and has some role in some of the boundless applications in the field of computer vision and image analysis with respect to machine learning strategies.

Image data is continuously increasing in both size and importance throughout the present time and in the near future. Working on image data is normally conducted using the image processing techniques, which are considered as necessary tools. Such algorithms are developed to simplify the work and make it much easier, while dealing with large amounts of data. Clustering techniques, for instance are based on the concept of breaking a set of data into some predefined subsets, such that the partition optimizes a given criterion. Clustering approaches are normally classified as being fuzzy or crisp (hard), according to the clustered subsets.

In hard clustering approaches, the data are divided up into a stated number of subsets, that are related such that each subset excludes the other. While, in fuzzy based clustering, the targeted items to cluster, may fit in to quite a lot of clusters with different degrees of membership.

With reference to the related published work, many researches in literature investigated with the Fuzzy C-Means (FCM) clustering algorithm in an extensive variety of ways aiming to achieve improved image segmentation results [1-3]. The FCM is an unsupervised method, which has been applied successfully to perform operations such as feature analysis [4], clustering [5-9], and classifier designs in fields such as astronomy [10], geology [11, 12], medical imaging, target recognition [13], and image segmentation.

In [14], Aljebory, and Mohammed presented an exhaustive investigation of the topic, where a technique for segmenting noisy medical images is proposed. The suggested modified fuzzy c-means clustering (MFCM) improves the Spatial Fuzzy c-means clustering (SFCM) algorithm performance. We would like to acknowledge that the work in our paper is an extension to the modified c-means clustering algorithm that was developed and presented in [14].

Fuzzy C-Means (FCM) assigns pixels to each class by means of fuzzy membership function. Assuming  $X=(x_1, x_2, x_3...x_N)$  represents an image with  $N$  pixels to be segmented into  $c$  clusters. FCM represents the minimization of the objective function expressed in (1) in an iterative manner:

$$CF = \lambda \times S_{ra}(c_i) + \frac{1-\lambda}{S_{er}(c_i)} \quad (1)$$

Where  $u_{ij}$  is the membership of pixel  $x_j$  in the  $i^{th}$  cluster,  $v_i$  is the  $i^{th}$  cluster center,  $m$  is the fuzzifier that controls the fuzziness of partitions and lies within the range ( $1 < m \leq \infty$ ), and  $\|\cdot\|$  is a norm metric. The center of each cluster may either be approximated or randomly initialized [15 -17].

Another famous method is the  $k$ -means clustering, in which a set of  $n$  objects is partitioned into  $k$  clusters so that the resulting intra-cluster similarity is high but the inter-cluster similarity is low. In  $k$ -means algorithm,  $k$  data values are assigned as initial cluster centers, and a set of  $n$  objects is divided up into  $k$  clusters such that the resulting inter-cluster similarity is high but the inter-cluster similarity is low. Eventually, the  $k$ -means algorithm minimizes an objective function  $J$ , which is expressed in (2).

$$Jk = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (2)$$

In equation (2), the symbol  $\|\cdot\|$  refers to a distance measure (intra) between a pixel  $x_i$  and the center of the cluster, it indicates the distance of cluster center from their  $n$  data points.  $K$ -means algorithm aims to minimize the sum of distances from each object to its cluster centroid. The algorithm iterates over all clusters and can be abstracted by the following steps:

1. Set  $k$  points into the space of the objects to be clustered. The  $k$  points represent initial  $k$  group centroids.
2. Based on minimum distance between the center and each pixel  $d = x_i - c_j$ , make assignment for objects

(pixels) and the closest centroid (assign pixels to the nearest center)

3. Recalculate the optimal new positions of the centers using the average relation and assign them the  $k$  centroids.
4. Repeat Step 2 and 3 until the centroids values stabilized with no change (i.e. it satisfies the tolerance or error value).

Despite their popularity in application, a disadvantage of both clustering methods (i.e.  $k$ -means, and  $c$ -means) is noticed. In both methods, it is not possible to find the appropriate number of clusters, and hence the application of the algorithms depends upon the clusters to be identified in advance [18]. In this paper, a new technique is proposed to estimate the optimal number of clusters by linking the  $k$ -means algorithm results with the  $c$ -means algorithm.

This paper structured as follows. Section 1 introduces the notions and importance of image processing and clustering. The concepts of the various image segmentation methods are also presented. In section 2, the FKM and FCM algorithms are compared, as a scientific foundation, to introduce the new algorithm. Section 3 presents the implementation of the proposed CFKCM algorithm together with an application example. Section 4 illustrates the experimental results to evaluate the performance of the proposed algorithm with respect to both FKM, and FCM techniques. In section 5, the analysis of the obtained results is presented. Finally, section 6 contains the conclusions and future work.

## 2. Fuzzy K-Means /C-Means Clustering

Clustering, as a term denotes the assignment of points to clusters such that in the same cluster they are as similar as possible, while points belonging to different clusters are not similar. In Fuzzy clustering, however, a data point can belong to more than one cluster.

As a well-known and widely used clustering algorithm, the standard FCM was primarily presented by Bezdek *et al.* [16], and its model was described by Ahmed *et al.* [19]. This algorithm is an iterative, unsupervised clustering algorithm, which can be used for image segmentation, where it considers the image as a dataset  $X$ ;  $X=\{x_1, x_2... x_N\}$ . Then the dataset  $X$  is subdivided into clusters that produce an optimal number of  $c$  partitions with centers vector  $V= \{v_1, v_2... v_c\}$  depending on memberships and satisfying the objective function in (1). So, the relation for  $v_i$  is deduced as per (3).

$$v_i = \frac{\sum_{j=1}^N u_{ij}^m x_k}{\sum_{i=1}^c u_{ij}^m} \quad (3)$$

The FCM algorithm iteratively optimizes  $Jc$  (given in (1)) by continuously updating the values of  $u$  and  $v$ , until the difference in consecutive  $u$  values is very small.

As a performance evaluation, FCM works efficiently for the data sets which are not corrupted with noise, while it is not reliable and may wrongly classify noisy data. Furthermore, the standard FCM computes the memberships based on the Euclidean distance between pixels. This affects the clustering performance especially when neighboring pixels are having strong correlation such as the medical MR images. While the  $k$ -Means clustering algorithm determines hard (crisp) clusters, where a point belongs to only one cluster, Fuzzy K-Means (FKM) discovers soft clusters, where a particular point can belong to more than one cluster, considering a certain likelihood. The functions involved when applying FKM are the mapping of a given set of representative vectors into an improved one through partitioning data points. The algorithm starts with a set of initial cluster centers and repeatedly performs the mapping process until a stopping criterion is satisfied, assuming that no two clusters are having the same cluster representative. If two cluster centers agree, one should be perturbed to avoid agreement in the iterative process. Similar to what we have introduced about the FCM algorithm, the FKM aims to optimize  $J_k$  in (4).

$$J_k = \sum_{j=1}^k \sum_{i=1}^n (u_{ij})^m (d_{ij})^2 \quad (4)$$

In equation (4),  $d$  denotes the Euclidean distance between data sample,  $V_i$  and the  $j^{\text{th}}$  center  $C_j$ .  $U_{ji}$  is the membership that determines the degree of belongingness of each data sample to all the centers.  $m$  on the other hand, determines the degree of fuzziness, when  $m=1$  the fuzzy clustering is reduced to the hard clustering algorithm.

In order to realize the motives behind the objective of this paper, a case of MRI image segmentation is considered; when applying the K-means, it is possible to detect a tumor or abnormality. While, by using C-means we can predict abnormality and tumor cells being either missed or not predicted by K-means algorithm. However, adopting C-means algorithm for clustering will reveal good results in segmenting noiseless images, but not with noisy images. Considering these facts, benefits may be obtained from integrating the two algorithms and having the number of iterations reduced, which affects the execution time and give an accurate detection. Both FCM and FKM algorithms perform the steps illustrated in the flowchart of Figure. 1.

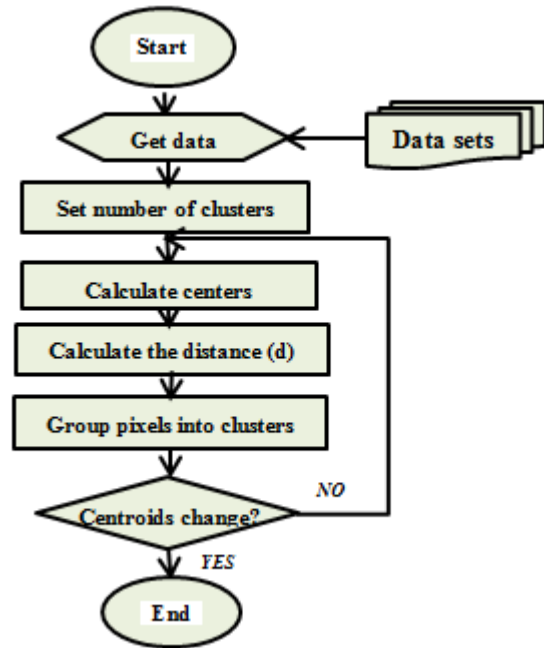


Figure 1. Algorithmic steps for clustering process

We have developed an algorithm that integrates both algorithms (FKM and KCM). The proposed approach, which we referred to as Combined Fuzzy k-means/c-means (CFKCM) algorithm, aims to emphasize the features and utilize them, while and overcoming the limitations of FKM and FCM techniques. The CFKCM approach, as a clustering algorithm, is used, and its performance is evaluated in the process of clustering MR medical images. The steps that the CFKCM algorithm follows are: pre-processing, clustering, object extraction and contouring, and validation stages.

By integrating the FKM and the FCM algorithms, the CFKCM is developed and the following two important drawbacks are targeted.

- Fuzzy C-means clustering techniques need to go through large number of iterations before reaching to the right cluster centers and having qualitative results level for clustering.
- Despite the certain fact that the Fuzzy K-means algorithm, will converge, optimal results may not be obtained.

The initial set of clusters and the number of clusters generated by the FCM clustering, are two issues on which the quality of the segmented image depends. To improve the clustering process, individual clusters should each have members with a strong relationship formed between them and the difference of belongingness among the clusters being minimized. The proposed algorithm performs the FKM clustering on the given dataset first. Resulted FKM centers are then used as cluster seeds, the FCM are performed on the given dataset, and clusters centers until reaching the termination conditions. The schematic diagram of the algorithm is as shown in Figure 2.

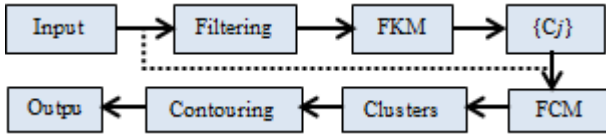


Figure 2. Block diagram of the proposed algorithm

It is illustrated in Figure 2 that the CFKCM algorithm starts by eliminating the noise from the MRI. The FKM algorithm is then adopted, where an initial value of clusters  $k$ , the total iterations limit, and the stopping criterion, are initialized. In the succeeding stage, the cluster centers  $\{C_j\}$  computed by FKM, are injected into the FCM algorithm to optimize its performance.

The second step is assigning each pixel value to the closest cluster focus based on a base distance between the pixel and the center of the cluster. Then the next stage starts by re-calculating the new centers by FCM. This continues until a pre-defined convergence rule is met. The memberships, distances and mean parameters are refreshed, then pixels in cluster and that scattered are reassigned according to the new parameters. Although, pixels are clustered once again by FCM depending on its altered cluster membership, there is no deduction by pixels in their groups.

With reference to the paper by Aljebory, and Mohammed [14], the membership function is defined by;

$$u_{ik} = \frac{\left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}}}{\left(\|x_k - v_i\|^{\frac{2}{m-1}}\right)} \quad (5)$$

The value of clustering fitness  $\lambda$  is calculated by substituting (5) in the following identity constraint

$$\sum_{j=1}^C u_{jk} = 1, \quad \forall k \quad (6)$$

So, we get;

$$\left(\frac{\lambda}{m}\right)^{\frac{1}{m-1}} = \frac{1}{\sum_{j=1}^C \left\{\|x_k - v_j\|^{\frac{2}{m-1}}\right\}^{-1}} \quad (7)$$

Where  $m \in (1, \alpha)$  is a weighting exponent which determines the fuzziness of the cluster. The main objective of the algorithm is to make best use and

maximize the intra-cluster similarity and minimize the inter-cluster similarity. In this work, the clustering fitness  $\lambda$  is considered as a performance measure, which depends on both similarities ( $0 < \lambda < 1$ ). The clustering fitness has a higher value when the inter-cluster similarity is low and vice-versa.

## 2.1. CFKCM Algorithm Flow

At the beginning, the proposed CFKCM clustering algorithm applies the FKM technique on a given dataset. Resulting cluster centers are then used as cluster seeds to the FCM.

Assuming an object  $X = \{x_1, x_2, \dots, x_n\}$ , where  $x_i$  is an element point. Then the fuzzy clustering is making a collection of  $k$  clusters,  $C_1, C_2, \dots, C_k$ , and a membership matrix  $U$  in domain  $[0, 1]$ , for  $i = 1 \dots n$  and  $j = 1, 2, \dots, k$ , where each element of the matrix  $U$  has a weight representing the degree of membership of object  $i$  in cluster  $C_j$ . The algorithm uses the notations;

$d_{ij}$  - The distance between  $i^{\text{th}}$  data and  $j^{\text{th}}$  cluster center.

$c$  - The number of clusters.

$m$  - The fuzziness index.

$U_{ij}$  - The membership for  $i^{\text{th}}$  data and  $j^{\text{th}}$  cluster center.

$n$  - The number of data points.

$c_j$  - The  $j^{\text{th}}$ -cluster center.

The condition to satisfy is that; the number of points in a cluster is at least one, and not all of the points, and the constraint given by:

$$\sum_{j=1}^k u_{ij} = 1, \quad \text{for } i=1 \text{ to } n \quad (8)$$

## 2.2. FKM Algorithm Steps

**Step 1:** From images dataset, randomly select  $k$  vectors as the initial centers,  $c_j$  ( $j=1, 2, \dots, k$ ).

**Step 2:** Use Euclidean distance  $d(x_i, c_j)$  to assign each data object to the closest cluster center.

Euclidean distance  $d(x_i, c_j) = \sqrt{\sum (x - c)^2}$ , for all  $i$  and  $j$  ( $j = 1, \dots, k$ ), where  $i = 1$  to  $N$ , and ( $j = 1, \dots, k$ ).

**Step 3:** Update cluster centers and compute the change. If change  $\geq \alpha$ , repeat steps 2-3.

The k-means (KM) algorithm is modified into fuzzy k-means (FKM) clustering algorithm as follows;

**Step 1:** Set the values for the initial cluster centers  $C_0 = \{C_j(0)\}$  and for  $\epsilon$ .

**Step 2:** Use Euclidean distance  $d(x_i, c_j)$  to assign each data object to the closest cluster center.

Euclidean distance  $d(x_i, c_j) = \sqrt{\sum (x - c)^2}$ , for all  $i$  and  $j$  ( $j = 1, \dots, k$ ), where  $i = 1$  to  $N$ , and ( $j = 1, \dots, k$ ).

**Step 3:** Based on the cluster centers update memberships  $U_{i,j}$  using the following equation (given  $m$  is the fuzzy index):

$$u_{ij} = \frac{1}{\left( (d_{ij})^{\frac{1}{m-1}} \sum_{n=1}^k \frac{1}{(d_{in})^{\frac{1}{m-1}}} \right)} \quad (9)$$

**Step 4:** Use the following formula to compute the new center for each cluster

$$c_j = \frac{\sum_{i=1}^n \frac{u_{ij}^m x_i}{u_{ij}^m}}{\sum_{i=1}^n \frac{u_{ij}^m}{u_{ij}^m}} \quad (10)$$

### 2.3. FCM-Algorithm Steps

**Step 1:** Parameters Initialization: set k vectors randomly as cluster means; set the initial membership matrix U, and also set the current iteration t = 0.

**Step 2:** Assign each data object X<sub>i</sub> to clusters using the membership matrix.

**Step 3:** Calculate the j<sup>th</sup> cluster center based on (11).

$$c_j^{t+1} = \frac{\sum_{i=1}^n \frac{u_{ij}^m x_i}{u_{ij}^m}}{\sum_{i=1}^n \frac{u_{ij}^m}{u_{ij}^m}} \quad (11)$$

**Step 4:** Use (12) to compute the new membership matrix,

$$U_{ji}^{t+1} = \left[ \sum_{l=1}^k \left( \frac{\|x_i - c_j\|}{\|x_i - c_l\|} \right)^{\frac{1}{m-1}} \right]^{-1} \quad (12)$$

**Step 5:** Based on the resulted membership matrix assign each data object X<sub>i</sub> to a cluster.

**Step 6:** Stop the process, if the numbers of vectors assigned to each cluster are not changing, if not, set t = t + 1 and repeat steps 3-5.

### 2.4. CFKCM-Algorithm Steps

In our current research, better clustering features are achieved by adopting the proposed concept of hybridization of FKM and FCM techniques. The CFKCM algorithm optimizes the segmentation algorithm by optimizing the initial cluster centers. The steps of the CFKCM algorithm are derived by combining the two algorithms that are based on fuzzy membership matrix. CFKCM first performs the FKM technique completely on the given dataset. The FCM will then adopt the given dataset based on the cluster centers resulted from the FKM as cluster seeds. This will continue until termination condition is reached. We emphasize that for running the first iteration of the FCM, the membership matrix, as well as the cluster centers are considered and calculated based on the results of FKM. Other iterations, that are remaining, continue as in the FCM algorithm.

The following steps describe the CFKCM algorithm for image segmentation;

**Step 1:** Set initial cluster centers C<sub>j</sub> (j = 1... k) randomly from dataset, and set t = 0.

**Step 2:** Start the FKM algorithm steps and continue processing through its steps.

**Step 3:** End processing if the stopping criterion is reached, else repeat step 2.

**Step 4:** Using the set of k vectors generated by FKM algorithm start the FCM clustering.

**Step 5:** Compute the new center and membership matrix. Use the membership matrix to allocate data objects to clusters, and continue processing through its steps.

**Step 6:** Compute the percentage of change in centers and stop if the change < η, η is a small constant. Or else, set t = t + 1, and repeat steps 3-5.

For all of the algorithms FKM, FCM, and CFKCM the same termination criterion is used with different values (α, ε, η). The CFKCM algorithm flowchart is illustrated in Figure 3.

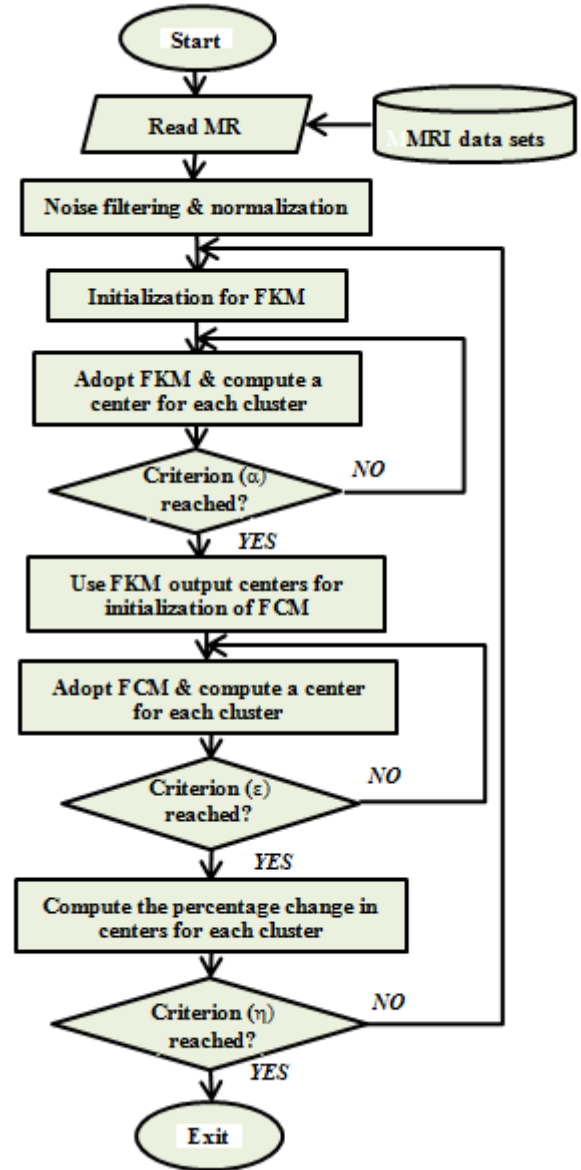


Figure 3. The flowchart for the CFKCM algorithm

## 3. CFKCM Algorithm Application

In this section, the application of the CFKCM



algorithm for spine MR images is presented. A total of 20 datasets for similar patients are used for testing the algorithm performance. Samples of the images are shown in Figure 4.

The clearest images in spinal cord and vertebra disks are selected for study. All MRIs are in DICOM format of dimensions (512 by 512) and (768 by 768) pixels. In order to study the consequence due to noise filtering on segmentation, and to assure the preprocessing importance, the MR image is subjected to a uniformly distributed random noise with variable magnitude before testing.

The spine consists of 33 individual bones stacked on each other. More than 120 MR images are taken for each case distributed by the main regions; neck (cervical), low back (lumbar), thoracic and sacral. These images presenting the anatomical structures on spine MRI (weighted T1, T2 axial and sagittal views) they are different in signal types; T1 weighted image (also referred to as T1WI or the "spin-lattice" relaxation time) is one of the basic pulse sequences in MRI and demonstrates differences in the T1 relaxation times of tissues [20]. The T2 relaxation refers to decline of transverse magnetization due to a combination of spin-spin relaxation and magnetic field inhomogeneity.

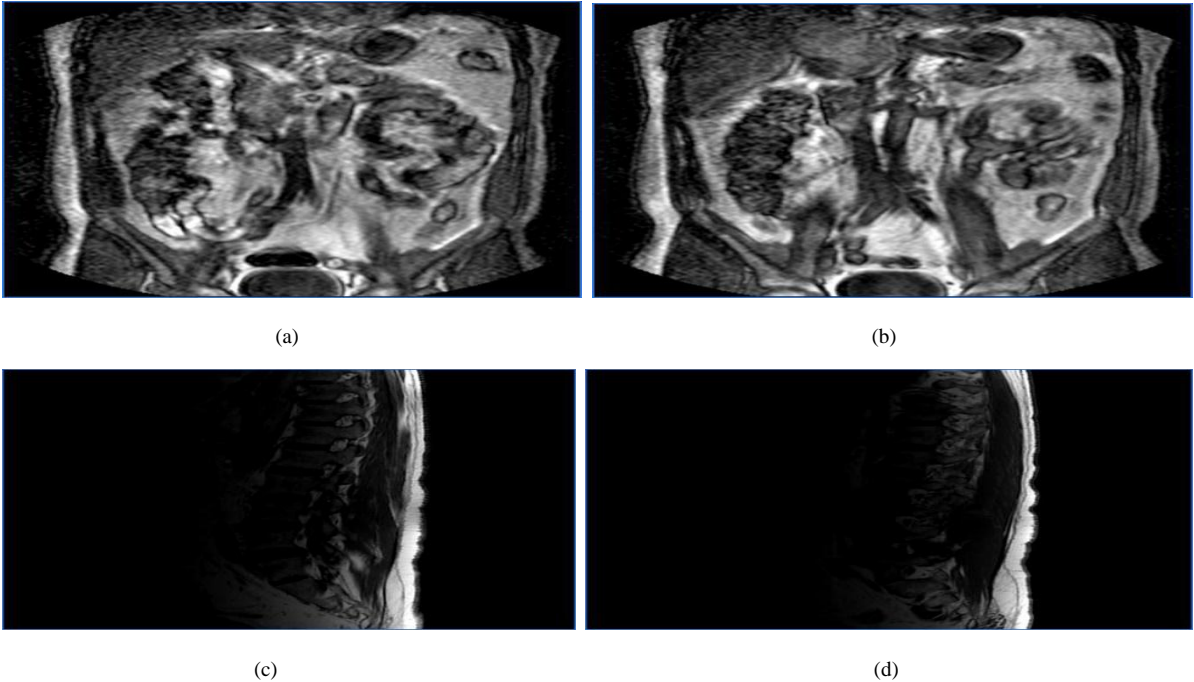
We focus on the applicability towards a large variety of clinical MR imaging sequences, dealing with low image quality, high anisotropy and spine pathologies. The algorithm can be switched into any one of the three modes; FKM, FCM, and the combined CFKCM mode. It was tested on segmenting spine section MRI images corrupted

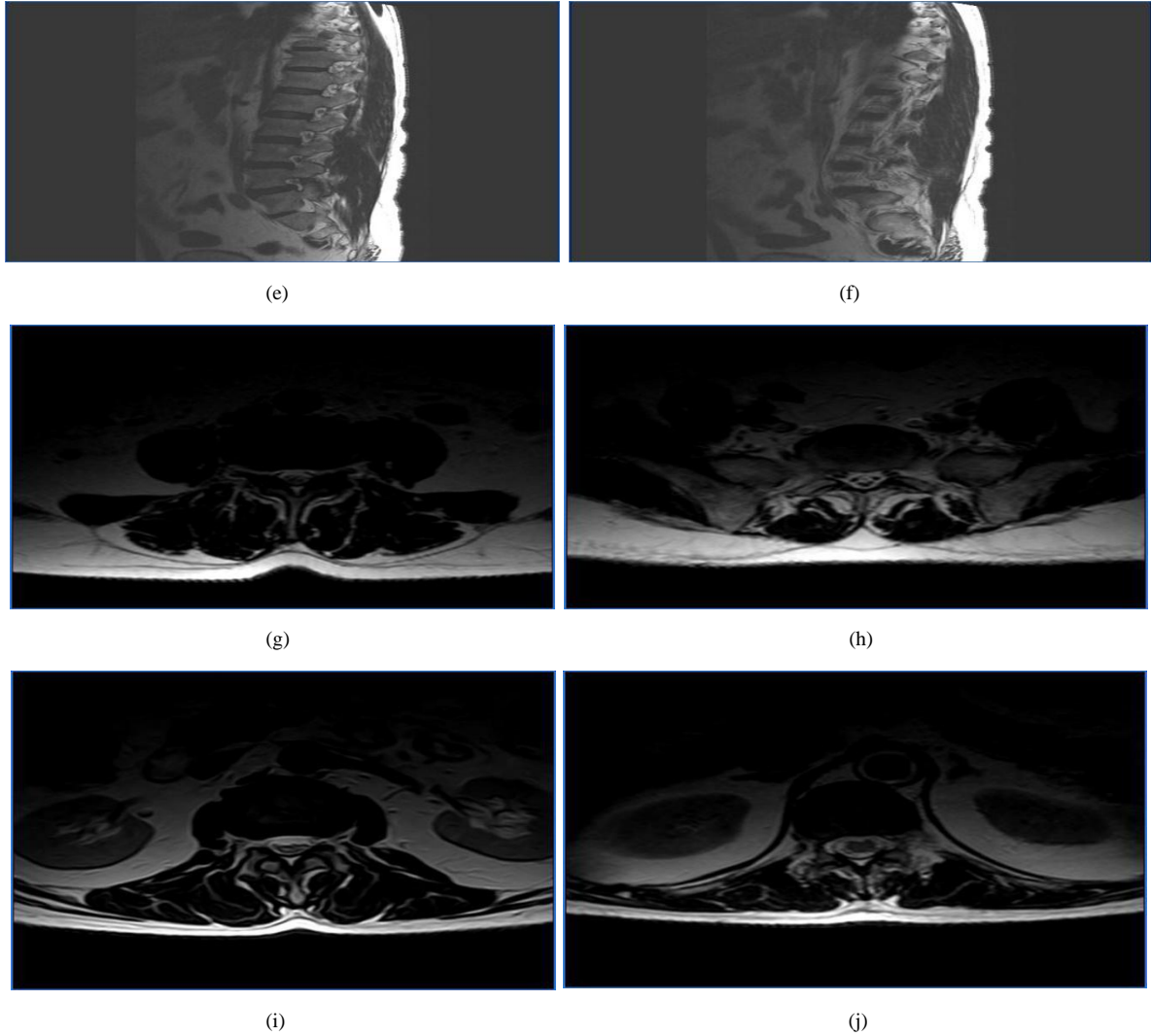
by different types of noise; zero centered symmetric  $\alpha$ -stable noise, white noise, salt and pepper noise, and mixed noise. In this experiment, the mixed noise used is the blend of Gaussian white noise  $N(0,100)$  and unit dispersion, zero centered symmetric  $\alpha$ -stable ( $S\alpha S$ ) noise. Variable amounts of noise and noise parameters are intentionally used in order to ensure the robustness of the algorithm. During experimentation, three different levels for the parameters controlling the intensity of the noise are adopted: Gaussian white noise (3%, 6%, and 9%), salt and pepper (4%, 8%, and 12%), while the parameter for the mixed noise  $\alpha$ -stable, is set to ( $\alpha=0.3, 0.6, \text{ and } 0.9$ ).

One more evaluation test of the performance of the three algorithms under different level of noises is also performed. This test is the Segmentation Accuracy (SA) test, in which the quality of the algorithms is measured by Segmentation Accuracy (SA) metric. SA is the percentage ratio of correctly classified pixels number to the total number of pixels, and is computed as per (13), by dividing the sum of pixels correctly assigned to a particular cluster  $A_i$  divided by the sum of the total number of pixels  $R_j$  in an image of 'c' clusters.

$$SA = \sum_i \frac{A_i \cap R_i}{\sum_j R_j} \quad (j \text{ and } i=1 \dots c) \quad (13)$$

The Segmentation Accuracy (SA) determines the fraction of correctly assigned pixels to a particular cluster, and reveals a clear idea about the noise removing capabilities of different algorithms. Higher value of SA is required for appropriate clustering.





**Figure 4.** A sample set of spine MR images.(a,b) SURVEY images, (c,d) T1W\_TSE\_sag images, (e,f) T2W\_TSE\_sag images, (g,h) eT2W\_TSE\_tra images, (i,j) T2W\_TSE\_tra images

### 3.1. CFKCM Algorithm Implementation

Let's consider the image  $I(R \times S)$  with  $R$  columns and  $S$  rows to be clustered,  $p(x,y)$  the pixel in consideration, where  $x = 1, 2, \dots, R$ , and  $y = 1, 2, \dots, S$ . Based on the analysis and algorithm steps, the position for each cluster center is calculated using (14):

$$c_j = \frac{1}{n_{c_j}} \sum_{x \in c_j} \sum_{y \in c_j} p(x,y) \quad (14)$$

Moreover, the assignment of every pixel to more than one cluster is found through calculating the membership function, according to equation (15):

$$U_{jp(x,y)}^m = \left[ \sum_{k=1}^{c_j} \left( \frac{d_{jp}}{d_{kp}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (15)$$

Where the distance from point  $(x,y)$  to the cluster  $j$  center is  $d_{jp}$  and  $d_{kp}$  is the distance from point  $p(x,y)$  to other cluster  $k$  center,  $c$  is the number of clusters and  $m$  is

the fuzziness index,  $m > 1$ , which regulate the degree of fuzziness. All pixels are assigned to a specific cluster depending on the belonging that based on the distance.

Normally, the closer the pixel, the maximum is the membership effect on the result of clustering process. The membership value is effectively based on amount of belonging. The normalized belongingness degree value  $M_j$  for every cluster is defined by;

$$M_j = \frac{c_j}{U_{jp(x,y)}^m} \quad (16)$$

High relation between the center and its pixels in the same cluster happens, when maximum degree of belongingness is achieved, thus we have the best data clustering results. In fact, the double refinement criterion in CFKCM algorithm (shown in Figure 3) helps each cluster to have pixels with a strong relationship formed between them, and the difference of belongingness between the clusters should be reduced [21].

## 4. Experimental Results

The proposed algorithm is implemented using Matlab version R2015a. Different types of medical images for spine are taken as source images and used for algorithm analysis and evaluation. In this section, we introduce results by which the performance of the three algorithms (FKM, FCM, and proposed CFKCM) is presented and compared. Several sets of real images have been used with the values of  $(\alpha, \epsilon, \eta)$  are set to be small values;

$$\alpha = 0.0001, \quad \epsilon = 0.00005, \quad \text{and } \eta = 0.00001$$

A main metric that is used as a performance comparison indicator of the algorithms under experimentation is the capability of the specific algorithm for noise removal. The algorithms (FKM and FCM) are tested to remove different types of noise (Gaussian, Salt and pepper, and mixed noise), while varying the noise level. It tests and assigns the pixels to proper clusters as indicated by the maximum values of SA.

### 4.1. Segmentation Accuracy (SA) Test

The SA metric, which is expressed in the formula of (13), is practically measured by conducting experiments on a dataset of 40 real MRIs. The MR images dataset is composed of cervical (neck) vertebrae numbered as C1 to C7, thoracic vertebrae numbered T1 to T12, lumbar vertebrae numbered L1 to L5, and four fused bones of the coccyx. An additive noise with different types and variable levels is added to measure the algorithm sensitivity performance as in the following four experiments.

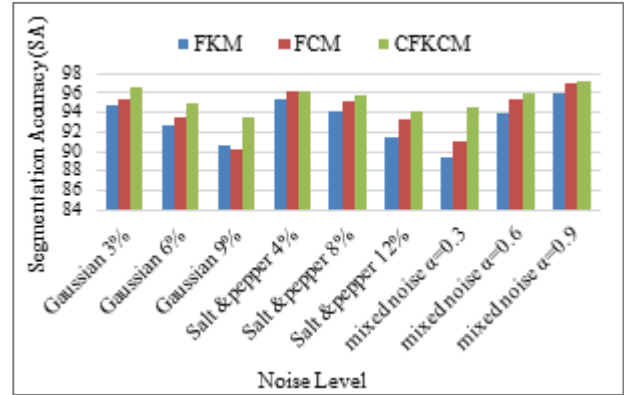
#### 4.1.1. Experiment 1

Test data: neck vertebrae C1-C7 MR images, dimensions (512 by 512).

**Table 1.** SA of the three algorithms on C1 to C7 MRI

Noise level	FKM	FCM	CFKCM
Gaussian 3%	94.75	95.35	96.56
Gaussian 6%	92.63	93.47	94.97
Gaussian 9%	90.56	90.23	93.48
Salt &pepper 4%	95.35	96.24	96.27
Salt &pepper 8%	94.10	95.26	95.73
Salt &pepper 12%	91.52	93.37	94.07
mixed noise $\alpha=0.3$	89.33	91.13	94.57
mixed noise $\alpha=0.6$	93.89	95.28	95.93
mixed noise $\alpha=0.9$	95.99	96.91	97.24

The Segmentation Accuracy (SA) values, tabulated in Table 1, are plotted in figure 5. From this figure, it's visually clear that the SA values for the CFKCM algorithm compared with the other algorithms are the highest.



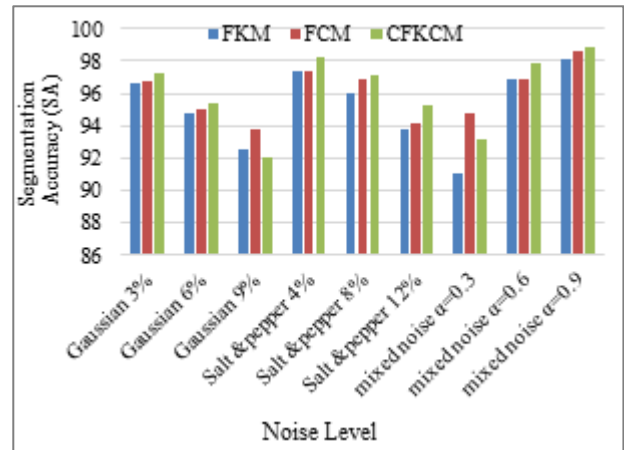
**Figure 5.** SA of the three algorithms on C1 to C7 MR images

#### 4.1.2. Experiment 2

Test data: thoracic vertebrae are numbered T1 to T12, dimensions (768 by 768). The calculated SA values are shown in Table 2, and plotted in Figure 6.

**Table 2.** SA of each algorithm on T1 to T12 images

Noise level	FKM	FCM	CFKCM
Gaussian 3%	96.65	96.75	97.28
Gaussian 6%	94.72	94.98	95.42
Gaussian 9%	92.53	93.77	92.03
Salt &pepper 4%	97.32	97.37	98.21
Salt &pepper 8%	96.02	96.89	97.14
Salt &pepper 12%	93.72	94.14	95.32
mixed noise $\alpha=0.3$	91.04	94.75	93.10
mixed noise $\alpha=0.6$	96.87	96.93	97.87
mixed noise $\alpha=0.9$	98.17	98.63	98.89



**Figure 6.** SA of each algorithm on T1 to T12 images

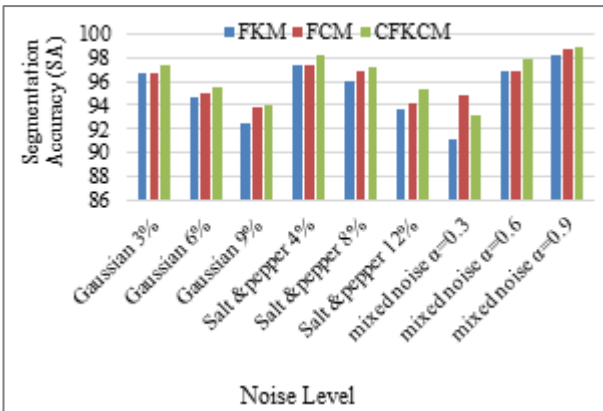
#### 4.1.3. Experiment 3

Test data: lumbar vertebrae numbered L1 to L5, dimension (768 by 768). For this set of images, SA values are calculated and presented in Table 3, and plotted in Figure 7.



**Table 3.** SA of the three algorithms on L1 to L5 MRI

Noise level	FKM	FCM	CFKCM
Gaussian 3%	96.65	96.75	97.28
Gaussian 6%	94.72	94.98	95.42
Gaussian 9%	92.53	93.77	94.03
Salt &pepper 4%	97.32	97.37	98.21
Salt &pepper 8%	96.02	96.89	97.14
Salt &pepper 12%	93.72	94.14	95.32
mixed noise $\alpha=0.3$	91.04	94.75	93.10
mixed noise $\alpha=0.6$	96.87	96.93	97.87
mixed noise $\alpha=0.9$	98.17	98.73	98.89



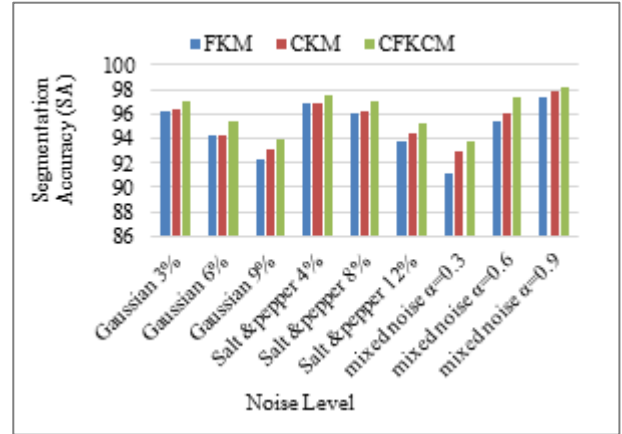
**Figure 7.** SA of the three algorithms on L1 to L5 MRI

4.1.4. Experiment 4

**Table 4.** SA of the three algorithms based on S1-S5 MR images

Noise level	FKM	FCM	CFKCM
Gaussian 3%	96.17	96.35	97.06
Gaussian 6%	94.22	94.19	95.32
Gaussian 9%	92.35	93.16	93.97
Salt &pepper 4%	96.85	96.89	97.46
Salt &pepper 8%	96.05	96.18	97.05
Salt &pepper 12%	93.76	94.37	95.21
mixed noise $\alpha=0.3$	91.15	92.86	93.69
mixed noise $\alpha=0.6$	95.37	96.12	97.36
mixed noise $\alpha=0.9$	97.28	97.89	98.17

Test data: four fused bones of the coccyx, dimensions (512 by 512). The SA values that are related to these images are shown in Table 4, while Figure 8, illustrating them graphically.



**Figure 8.** SA of the three algorithms on S1-S5 MR images

4.2. Clustering Fitness Test

A large number of clusters lead to the involvement of more details which can eventually help in minute object detection. To show this, we choose a diversely illuminated images ‘neck, mid back, low back, and tailbone’ from the real dataset, which contains differently illuminated layers. All the algorithms; FKM, FCM, and the proposed CFKCM are executed on each image data with variable number of clusters ( $c = 6, 7, 8, 9, 10, 11, 12, 13, 14$ ). For all algorithms, the same cluster seed and termination conditions are used. The Clustering Fitness (CF) for one of the possible clustering results  $C$ , is calculated according to (17) [22].

$$CF = \lambda \times S_{ra}(c_i) + \frac{1-\lambda}{Ser(c_i)} \quad (17)$$

Where  $\lambda$  is the experiential weight, ( $0 < \lambda < 1$ ),  $S_{ra}(c)$  is the intra-cluster and  $Ser(c)$  is the inter-cluster similarity for the cluster  $c_i$ . To avoid any bias,  $\lambda$  is supposed to be 0.5. The CF for each algorithm is tabulated in Table 5.

**Table 5. (a-d)** The CF value for the three algorithms

c	FKM	FCM	CFKCM
6	29.75	29.35	33.87
7	29.10	29.00	33.25
8	28.25	28.55	32.75
9	27.03	28.45	32.58
10	26.75	27.05	32.51
11	26.95	27.83	33.50
12	28.74	29.17	33.76
13	30.82	30.05	33.95
14	31.25	30.65	34.25

(a)

c	FKM	FCM	CFKCM
6	27.35	28.05	30.25
7	27.04	27.36	29.37
8	26.12	27.00	28.25
9	25.35	26.15	27.85
10	24.82	25.85	27.15
11	24.75	25.80	26.55
12	25.47	26.55	27.20
13	26.30	27.63	28.32
14	28.15	29.15	30.45

(b)

c	FKM	FCM	CFKCM
6	28.45	28.95	33.20
7	27.13	28.03	32.67
8	26.38	27.34	32.45
9	25.79	26.74	32.24
10	25.57	26.50	32.15
11	26.85	27.89	33.27
12	28.23	29.35	33.68
13	30.15	30.52	33.82
14	30.78	30.86	33.95

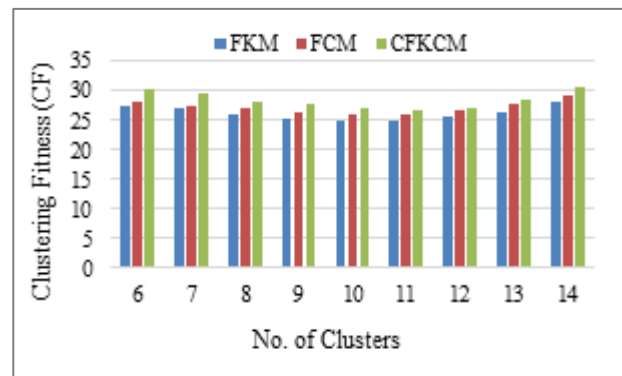
(c)

c	FKM	FCM	CFKCM
6	32.25	32.95	33.85
7	31.42	32.20	33.25
8	30.58	31.48	32.89
9	29.50	31.15	32.50
10	28.35	30.57	32.50
11	28.15	30.35	33.75
12	28.58	30.90	33.95
13	30.05	31.74	34.20
14	31.75	32.55	34.48

(d)

In Tables (5a- 5d) are the Clustering Fitness values (CF) being calculated for the four adopted categories of images (i.e. the cervical (neck) vertebrae, the thoracic vertebrae, the lumbar vertebrae, and the fused bones of the coccyx) respectively. In these results, we can notice that the CF value is almost constant and not related to the increase of the number clusters. Further analysis of these result, illustrates that the proposed CFKCM algorithm outperforms both other two algorithms. With CFKCM, that is having the highest CF value, it will simply enhance the validity of segmentation performance.

The results for the Vertebrae T1-T12 MR images are elected to visually illustrate the behavior of the three clustering algorithm in terms of the clustering fitness, and the values given in Table 5-b are plotted in Figure 9.



**Figure 9.** Cluster Fitness for FKM, FCM and CFKCM algorithms (Vertebrae T1-T12 MR images).

### 5. Results and Analysis

The experimental results show that the CFKCM enhanced algorithmic segmentation in the following attributes; a- the starting number of clusters is almost optimal, b- higher SA even with noisy images, c- robust algorithm that, maintains the structural characteristics of the segmented image, d- higher CF, where the proposed algorithm shows appreciable performance for all sorts of noises, e- the center of the cluster is located in the optimal gravity center of all elements that belongs to the cluster.

The performance of clustering is evaluated by applying fuzzy partition method. The partition coefficient  $V_{pc}$  and entropy  $V_{pe}$  are defined as follows [23, 24]:

$$V_{pc} = \frac{1}{n} \sum_{i=1}^c \sum_{k=1}^n u_{ik}^2 \tag{18}$$

$$V_{pe} = \frac{1}{n} \sum_{i=1}^c \sum_{k=1}^n u_{ik} \log u_{ik} \tag{19}$$

When  $V_{pc}$  maximum or  $V_{pe}$  minimum, the best performance is achieved. This test function is applied to the four types of images corrupted by different levels of Gaussian noise (5%, 9%, and 13%). The results for CFKCM and FCM are depicted as in Table 6.

For further illustration, part of the results is plotted in Figure 10, where the obtained Cluster validity function values for noise level = 13%, are emphasizing the expectation and telling that the performance of the CFKCM algorithm beats that of the FCM.

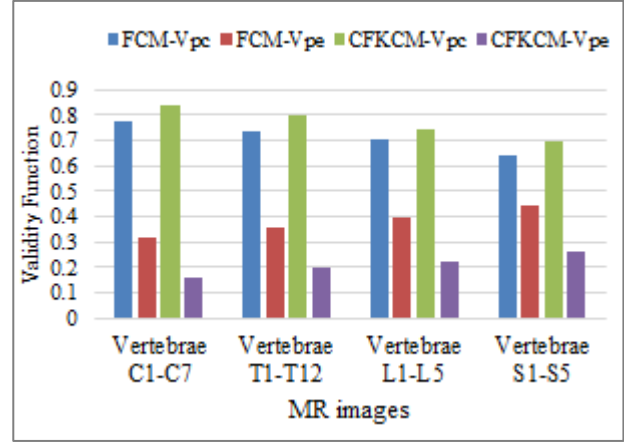


Figure 10. Cluster validity function for FCM and CFKCM algorithms (Noise level = 13%).

The FCM algorithm is sensitive to noise; noisy pixels are almost wrongly classified. This is because the objective function does not consider the spatial information, but deal with images as the same as separate points. The integration of more than one constraint by combined in CFKCM enhances the validity of segmentation performance. This can be seen, clearly, in the visual results for both the FCM and the CFKCM algorithm.

Table 6. Cluster validity function for FCM and CFKCM algorithms

Image	Noise Level	FCM		CFKCM	
		$V_{pc}$	$V_{pe}$	$V_{pc}$	$V_{pe}$
vertebrae C1-C7	5%	0.8225	0.2873	0.9275	0.0987
	9%	0.8023	0.2980	0.8723	0.1079
	13%	0.7725	0.3205	0.8357	0.1574
vertebrae T1-T12	5%	0.8073	0.2945	0.8923	0.1065
	9%	0.7911	0.3060	0.8650	0.1273
	13%	0.7390	0.3592	0.7965	0.2014
vertebrae L1-L5	5%	0.7860	0.3235	0.8105	0.1207
	9%	0.7219	0.3790	0.7860	0.2032
	13%	0.7010	0.3997	0.7468	0.2253
vertebrae S1-S5	5%	0.7695	0.3545	0.8025	0.1255
	9%	0.7105	0.3902	0.7495	0.2135
	13%	0.6429	0.4467	0.6953	0.2610

## 6. Conclusions

The paper presented a study of merging two of the popular clustering algorithms, namely the fuzzy k-means (FKM) and fuzzy c-means (FCM) and identifying a proposed strategy out of this combination process. The resulted algorithm is referred to as a Combined Fuzzy K-means, C-Means (CFKCM). The application steps of the CFKCM algorithm were explained and the performance, in terms of image clustering, has been studied. A dataset of MR images was used for experimentation and performance evaluation. The study is based on comparing the modified algorithm with the two other algorithms (i.e. fuzzy k-means, and fuzzy c-means). We took into consideration a number of factors that are adopted in the field for performance measurement. Four distinct experiments were carried out for segmenting a dataset composed of 40 MR images, and the results of clustering backbone MR images including seven cervical (neck) vertebrae, twelve thoracic vertebrae, five lumbar vertebrae, four fused bones of the coccyx, were obtained. Experimentation results showed that the proposed algorithm outperforms the other two algorithms based on the values of the Segmentation Accuracy (SA) as measuring metric. During the experiments an additive noise with different types, and variable levels is added to measure the algorithm sensitivity performance as in the following four experiments. The results showed that the noise effect with the MFCM was significantly less than with other algorithms. The clustering Fitness (CF) was also used as performance measurement metric, and CF tests were performed. Again, the three algorithms were tested, and the values of CF were calculated for MR images from the dataset with varying number of clusters ( $c = 6, 7, 8, 9, 10, 11, 12, 13, 14$ ). The same cluster seed and termination conditions are used for all tests. Calculations resulted in the CFKCM algorithm having the highest CF value, and this result is an indicator to an enhanced validity of segmentation performance. A third test for performance evaluation was the calculation of clustering validity functions (terms  $V_{pc}$  and  $V_{pe}$ ). With higher values of  $V_{pc}$ , and lower values of  $V_{pe}$ , the CFKCM algorithm exhibits better performance than FCM.

---

## REFERENCES

- [1] K. Kaltri and M. Mahjoub. Image segmentation by gaussian mixture models and modified FCM algorithm. The International Arab Journal of Information Technology. Vol.11, No., 11-18, 2014.
- [2] K. M. Aljebory, and Thabit. S. Mohammed, Developing Modified Fuzzy C-Means Clustering Algorithm for Image Segmentation, The International Multi-Conference on Systems, Signals and Devices 2018,(IEEE SSD'18), Hammamet, Tunis, 19-22, March 2018.
- [3] S. Wang, Z. Geng, J. Zhang, Y. Chen and J. Wang. A fuzzy C-means model based on the spatial structural information for brain MRI segmentation. International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol.7, No. 1, 313-322, 2014.
- [4] S. Xu, X. Liu, C. Li and L. Hu. An improved fast FCM image segmentation algorithm based on region feature analysis, Pattern Recognition and Artificial Intelligence, Vol. 25, No.6, 987-995, 2012.
- [5] W. Wiharto and E. Suryani, The Comparison of Clustering Algorithms K-Means and Fuzzy C-Means for Segmentation Retinal Blood Vessels, Acta Inform Med, Vol. 28, No.1, 42-47, 2020.
- [6] H. Shamsi and H. Seyedarabi, A Modified Fuzzy C-Means Clustering with Spatial Information for Image Segmentation, International Journal of Computer Theory and Engineering, Vol. 4, No. 5, 62-766, 2012.
- [7] K. Chuang, H. Tzeng, S. Chen, J. Wu, and T.-J. Chen, Fuzzy c-means clustering with spatial information for image segmentation, Computerized Medical Imaging and Graphics vol. 30, no. 1, 9-15, 2006.
- [8] M. Gong, Y. Liang, J. Shi, W. Ma, and J. Ma, Fuzzy c-means clustering with local information and kernel metric for image segmentation, IEEE Transactions on Image Processing, vol. 22, no. 2, 573-584, 2013.
- [9] X. Wang, and J. Bu, A fast and robust image segmentation using FCM with spatial information, Elsevier, 2009.
- [10] R. Hussain, Geological Featuring of Charon images using Fuzzy c-means (FCM) Clustering Algorithm, J. Appl. Environ. Biol. Sci., vol. 6, no. 6, 89-96, 2016.
- [11] G. Efendiyev, P. Mammadov, I. Piriverdiyev and V. Mammadov, Clustering of Geological Objects Using FCM-algorithm and Evaluation of the Rate of Lost Circulation, Procedia Computer Science, Vol. 102, 159-162, 2016.
- [12] A. Gary, M. Wakefield, G. Johnson, and D. Ekart, Application of Fuzzy C-Means Clustering to Paleoenvironmental Analysis: Example from the Jurassic, Central North Sea, UK, book chapter, ISBN 978-1-56576-137-7, 9-20, published January 01, 2009.
- [13] Y. Lei, S. Liu, and W. Kong, Target Recognition Based on Feature Weighted Intuitionistic FCM, Computer Vision. CCCV 2015. Communications in Computer and Information Science, vol. 546. Springer, Berlin, 2015.
- [14] K. M. Aljebory, and Thabit S. Mohammed, Modified Fuzzy C-Means Clustering Algorithm Application in Medical Image Segmentation, JEA Journal of Electrical Engineering, Vol. 2, No.1, 2018.
- [15] M. Sezgin, and B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, Journal of Electronic Imaging, Vol.13, No.1, 146-165, 2004.
- [16] J. Bezdek, L. Hall, and L. Clarke, Review of MR image segmentation using pattern recognition, Med Phys, Vol.20, 1033-48, 1993.
- [17] S. Le, S. Chung, and R. Park, A comparative performance study of several global thresholding techniques for segmentation, Graph. Models Image Process, Vol. 2, 171-

- 190, 1990.
- [18] D. Tsai, and C. Lin, Fuzzy C-means based clustering for linearly and nonlinearly separable data, *Pattern Recognit.*, Vol. 44, No. 8, 1750–1760, 2011.
- [19] M. Ahmed, S. Yamany, N. Mohamed, A. Farag, and T. Moriarty, Modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data, *IEEE Trans Med Imaging*, Vol. 21, 193–199, 2002.
- [20] R. Bitar, G. Leung, R. Perng, S. Tadros, A. Moody, J. Sarrazin, C. McGregor, M. Christakis, S. Symons, Andrew Nelson, and T. Roberts, MR Pulse Sequences: What Every Radiologist Wants to Know but is Afraid to Ask, *Radio Graphics*, Vol. 26, No. 2, 513-538, 2006.
- [21] X. Yang, and G. Zhang, A kernel fuzzy C-means clustering-based fuzzy support vector machine algorithm for classification problems with outliers or noises, *IEEE Trans. Fuzzy Syst.*, Vol. 19, No. 1, 105–115, 2011.
- [22] J. Raitoharju, K. Samiee, S. Kiranyaz, M. Gabbouj, “Particle Swarm Clustering Fitness Evaluation with Computational Centroids”, *Swarm and Evolutionary Computation*, Vol. 34, 103-118, 2017.
- [23] M. Sayadi, L. Tlig and F. Fnaiech, A New Texture Segmentation Method Based on The Fuzzy C-Mean Algorithm and Statistical Features, *Applied Mathematical Sciences*, Vol. 1, 2999 – 3007, 2007.
- [24] R. Shang, P. Tian, R. Stolkin, J. Feng, B. Hou and X. Zhang, A Spatial Fuzzy Clustering Algorithm With Kernel Metric Based On Immune Clone For Sar Image Segmentation, *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, Vol. 9, No. 4, 1640-1652, 2016.