We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Incorporating Local Data and KL Membership Divergence into Hard C-Means Clustering for Fuzzy and Noise-Robust Data Segmentation

Reda R. Gharieb

Abstract

Hard C-means (HCM) and fuzzy C-means (FCM) algorithms are among the most popular ones for data clustering including image data. The HCM algorithm offers each data entity with a cluster membership of 0 or 1. This implies that the entity will be assigned to only one cluster. On the contrary, the FCM algorithm provides an entity with a membership value between 0 and 1, which means that the entity may belong to all clusters but with different membership values. The main disadvantage of both HCM and FCM algorithms is that they cluster an entity based on only its self-features and do not incorporate the influence of the entity's neighborhoods, which makes clustering prone to additive noise. In this chapter, Kullback-Leibler (KL) membership divergence is incorporated into the HCM for image data clustering. This HCM-KL-based clustering algorithm provides two-fold advantage. The first one is that it offers a fuzzification approach to the HCM clustering algorithm. The second one is that by incorporating a local spatial membership function into the HCM objective function, additive noise can be tolerated. Also spatial data is incorporated for more noise-robust clustering.

Keywords: data science, clustering, image clustering, hard and fuzzy C-means, membership function, Kullback-Leibler (KL) divergence

1. Introduction

Image segmentation is a principle process in many image, video, scene analysis and computer vision applications [1–3]. The objective of segmentation process is to divide an image into multiple separate regions or clusters which make it easier to recognize and distinguish different objects in image. Over the last few decades, several image segmentation methods have

been developed. However, there is still no satisfactory performance especially in noisy images. This makes development of segmentation algorithms that are capable of handling noisy images an active area of research. The current segmentation methods can be classified into thresholding, region-detection, edge-detection, probabilistic and artificial neural-network classification and clustering [1–3]. Among the widely used are the hard and fuzzy-based clustering methods since clustering needs no training examples [4–24]. Hard C-means (HCM) also called K-means clustering algorithm is an unsupervised approach in which data is basically partitioned based on locations and distances between various data points [4–6]. K-means partitions the data into C-clusters so that the distances between data within each cluster are as close as possible but as far as possible between data in different clusters. HCM clustering algorithm offers crisp segmentation in which each data point belongs to only one cluster. Thereby it does not take into consideration fine details of infrastructure of data such as hybridization or mixing. Compared with HCM algorithm, fuzzy C-means (FCM) algorithm is able to provide soft segmentation by incorporating membership of belonging described by a membership function [7, 8]. However, one disadvantage of the standard FCM is not incorporating any spatial or local information in image context, making it very sensitive to additive noise and other imaging artifacts. To handle this problem, different techniques have been developed [9–13]. These techniques have involved spatial or local data information for the enhancement and regularization of the performance of the standard FCM algorithm. Local membership information has also been employed to generate a parameter to weight or modify the membership function in order to give more weight to the pixel membership if the immediate neighborhood pixels are of the same cluster [14]. HCM algorithm has also been fuzzified by involving membership entropy optimization [15–17].

In this chapter, HCM clustering algorithm is modified by incorporating local spatial data and Kullback-Leibler (KL) membership divergence [18–22]. The local data information is incorporated via an additional weighted HCM function in which the smoothed image data is used for the distance computation. The KL membership divergence aims at minimizing the information distance between the membership function of each pixel and the locally smoothed one in the pixel vicinity. The KL membership divergence thus provides an approach for regularization and fuzzification. The chapter is organized as follows. In Section 2, clustering problem formulation is overviewed. In Section 3, HCM clustering algorithm is described. In Section 4, several FCM-related clustering algorithms are explained. In Sections 5 and 6, the proposed local membership KL divergence-based FCM (LMKLFCM) and Local Data and membership KL divergence-based FCM (LDMKLFCM) clustering algorithm are discussed. In Section 7, simulation results of clustering and segmentation of synthetic and real-world images are presented. Finally Section 8 presents the conclusion.

2. Problem formulation

The objective is to cluster a set of observed data $\{x_n; n = 1, 2, .., N\}$ where each data point is an $M-$ dimensional real-vector called the feature or the pattern vector, i.e., $x_n \in \mathrm{R}^{1 \times M}$. For gray-scale image data, $\{x_n; n = 1, 2, .., N\}$ is a row-wise concatenation of a 2-D image

$\{X_{pq}; p = 1, 2, .., P; q = 1, 2, .., Q\}$. That is $n = (p-1)Q + q$ and the intensity-feature x_n is a single-dimensional real-value, i.e., $M = 1$. Clustering aims at partitioning theses N observations into $C < N$ divisions, $\{\mu_1, \mu_2, …, \mu_C\}$ called C clusters or segments so as to make the entities or pixels in the same cluster as similar as possible and the ones in different clusters as dissimilar as possible. One approach to cluster these data is to minimize the within-clusters sum of squares of distances (WCSS) and to maximize the between-clusters sum of squares of distances (BCSS).

3. Hard C-means (HCM)

In hard C-means (HCM) algorithm also called the K-means one, the objective is to minimize the following function [4–6, 15].

$$J_{HCM} = \sum_{i=1}^{C} \sum_{n=1}^{N} u_{in} d_{in} \qquad (1)$$

where $d_{in} = \|x_n - v_i\|^2$, is the square of the Euclidian distance between the nth pixel feature x_n of the image under segmentation and $v_i \in V = \{v_1, v_2, …, v_C\}$ called the center of the ith cluster given by

$$v_i = \frac{\sum_{x_n \in \mu_i} x_n}{N_i}, \qquad i = 1, 2, …, C. \qquad (2)$$

where μ_i is the ith cluster label and N_i is its number of patterns in cluster i. In (2), it is clear that the pattern x_n belongs to only one cluster which means that $u_{in} \in \{0, 1\}$ called the membership function is given by [15].

$$u_{kn} = \begin{cases} 1; & k = \arg \min_i(d_{in}) \\ 0, & Otherwise \end{cases} \qquad (3)$$

From (3), it is obvious that the HCM provides a crisp membership function $u_{in} \in \{0, 1\}$ or {False, True}. $u_{in} \in \{0, 1\}$. Thus HCM algorithm does not take into account fine details of infrastructure

---

Given $x_n, \; n = 1, 2, …, N$.

Initialize $v_i^0, \; i = 1, 2, .., C; t = 0$;

1. For $n = 1, 2, …, N$
   Compute:
2. $d_{in} = \|x_n - v_i^t\|^2; i = 1, 2, …, C$.
3. $k = \arg \min_i(d_{in}); u_{kn}=1; u_{in} = 0; i = 1, 2, .., C; i \neq k$. (HCM);
   $u_{in} = \dfrac{1}{\sum_{j=1}^{C} \left(\frac{d_{in}}{d_{jn}}\right)^{\frac{1}{(m-1)}}}$ (FCM)
4. Update $t = t + 1; v_i^{t+1} = \dfrac{\sum_n u_{in} x_n}{\sum_n u_{in}}, \qquad i = 1, 2, …, C$.
5. Check if $\|V^t - V^{t+1}\|^2 > \varepsilon$ (negligible change); repeat 1–5 until convergence.

Table 1. Pseudo code of the HCM (FCM) algorithms.

such as hybridization or mixing of data which is important in data clustering and decision making. The algorithm is implemented by an iterative procedure as summarized in **Table 1**.

4. Related fuzzy clustering algorithms

4.1. Conventional FCM

The fuzzy C-means (FCM) algorithm seeks to minimize the following objective function [7].

$$J_{FCM} = \sum_{i=1}^{C} \sum_{n=1}^{N} u_{in}^m d_{in} \tag{4}$$

It is obvious that the difference between the FCM algorithm and HCM one is the incorporation of the exponent parameter m, called the fuzzification parameter, and if it is settled to 1, the FCM algorithm reduces to the HCM one. Thus, due to this exponent m, the membership of the nth pixel to the ith cluster, u_{in}, can take on an infinite set of values from 0 to 1. Thus each nth pixel may belong to all clusters with equal membership values of $1/C$ which in this case we obtain too fuzzy membership function. Then the exponent parameter $1 < m$ is incorporated to control the degrees of fuzzification; the bigger the m, the more the fuzzification. Finally, the fuzzy membership u_{in} should satisfy [7].

$$u_{in} \in U = \left\{ u_{in} \in [0,1]; \sum_{i=1}^{C} u_{in} = 1 \forall n; 0 < \sum_{n=1}^{N} u_{in} < N \forall i \right\}, \tag{5}$$

The membership u_{in} and the cluster-center v_i that minimize the FCM function in (4), subject to $\sum_{i=1}^{C} u_{in} = 1 \forall n$ are given by [7].

$$u_{in} = \frac{1}{\sum_{j=1}^{C} \left(\frac{d_{in}}{d_{jn}} \right)^{\frac{1}{(m-1)}}} \tag{6}$$

$$v_i = \frac{\sum_{n=1}^{N} u_{in} x_n}{\sum_{n=1}^{N} u_{in}} \tag{7}$$

4.2. Local spatial data-based FCM (LDFCM)

The neighboring pixels of an image are highly correlated and are thus highly expected to belong to the same cluster or object. To get benefit from this spatial data information, the standard FCM objective function in (4) has been modified by adding a weighted regularization function dependent on local image data information [10–12]. That is, the LDFCM objective function is given by

$$J_{LDFCM} = J_{FCM} + \alpha \sum_{i=1}^{C} \sum_{n=1}^{N} u_{in}^m \overline{d_{in}} \tag{8}$$

where α is a weight to be experimentally selected by the user, m is a fuzzification parameter, $\overline{d}_{in} = \| \overline{x}_n - v_i \|^2, \overline{x}_n \in \overline{X}$ is the nth pixel of the locally-smoothed image, $\overline{X}$, obtained in advance

from the original one by $\overline{X} = w^{(x)**}X$, where ** means two-dimensional convolution. The weights $w^{(x)}$ can be equal or not provided that its centerweight is zero and are summed to unity. From (8), it is clear that the LDFCM aims at minimizing the standard FCM objective function plus another weighted modified FCM function acting as a regularization function. In this regularization FCM function, the distances are generated from the locally-smoothed image data instead of the original image data. Therefore, this correlates the clustering pixel x_n with its immediate spatial neighboring pixels which biases the algorithm to provide clustered images with piecewise homogenous regions. The membership u_{in} and the cluster-center v_i functions of the LDFCM method are given by [10–12].

$$u_{in} = \frac{1}{\sum_{j=1}^{C} \left(\frac{d_{in} + \alpha \overline{d}_{in}}{d_{jn} + \alpha \overline{d}_{jn}} \right)^{\frac{1}{(m-1)}}} \tag{9}$$

$$v_i = \frac{\sum_{n=1}^{N} u_{in}(x_n + \alpha \overline{x}_n)}{(1 + \alpha) \sum_{n=1}^{N} u_{in}} \tag{10}$$

It is obvious from (9) and (10) that when $\alpha = 0$, the membership u_{in} and the cluster-center v_i become the ones provided by the standard FCM algorithm in (6) and (7). The advantage of the LDFCM method arises from involving the locally-smoothed data $\alpha \overline{x}_n$ in computing the membership u_{in} and the cluster-center v_i functions which indeed can handle additive noise.

4.3. Spatial-based fuzzy C-means (SFCM)

An approach to incorporating local spatial data information into the standard FCM has been presented in [13]. The objective function of the SFCM algorithm is given by

$$J_{SFCM} = \sum_{i=1}^{C} \sum_{n=1}^{N} u_{in}^m D_{in}, \tag{11}$$

where D_{in} is a modified or weighted distance between the nth pixel and the ith cluster-center. This modified distance is computed from the original or standard distance $d_{in} = \|x_n - v_i\|^2$ as follows

$$D_{in} = (1 - \lambda)d_{in}f_{in} + \lambda d_{in} \tag{12}$$

where $\lambda \in [0, 1]$ is an experimentally selected weight, and f_{in} is a spatial or local data function given by [13].

$$f_{in} = \frac{\sum_{k \in N_n} d_{ik}}{\min\left\{ \sum_{k \in N_n} d_{ik}; i = 1, 2, .., C \right\}} \tag{13}$$

It is obvious from (12) that with $\lambda = 1$, the SFCM clustering method reduces to the standard FCM method. The spatial data function f_{in} is dependent on the original distances of the set of pixels N_n in the immediate neighborhood of the nth pixel. If all pixels in the neighbor set do

not belong to the ith cluster f_{in} is maximum since the denominator is minimum while the numerator is maximum. This implies that f_{in} causes D_{in} to increase when the pixels of the immediate neighborhood of the nth pixel do not belong to the ith cluster. This increase of D_{in} contributes to decreasing the membership u_{in} for achieving and preserving the minim of the SFCM function in (11).

The membership u_{in} and the cluster-center v_i associated with the SFCM method are given by [13].

$$u_{in} = \frac{1}{\sum_{j=1}^{C} \left(\frac{D_{in}}{D_{jn}}\right)^{\frac{1}{(m-1)}}} \tag{14}$$

$$v_i = \frac{\sum_{n=1}^{N} u_{in} x_n}{\sum_{n=1}^{N} u_{in}} \tag{15}$$

It is obvious from (14) that similar to the standard FCM, the membership u_{in} is inversely proportional to the weighted distance D_{in}, which again means that, increasing D_{in} when the immediate neighboring pixels to the nth pixel do not belong to the ith cluster, decreases the membership function u_{in}. From (15), however, it is clear that the SFCM algorithm computes the cluster-center v_i in a similar way as the standard FCM method does. Hence, additive noise can still reduce the accuracy of cluster center v_i obtained by the SFCM algorithm.

4.4. HCM incorporating membership entropy

The membership entropy has been incorporated into the HCM for fuzzification. The membership entropy-based FCM (MEFCM) algorithm has the following objective function [17].

$$J_{MEFCM} = J_{HCM} + \beta \sum_{i=1}^{C} \sum_{n=1}^{N} \left(u_{in} \log\left(u_{in}\right) + (1 - u_{in}) \log\left(1 - u_{in}\right)\right) \tag{16}$$

where $\beta > 0$ is a weight experimentally selected to control the fuzziness of the entropy term. We still need U to be constrained to satisfy (5). It can be shown that the membership and the cluster-center that minimize (16) are respectively given by [17]

$$u_{in} = \frac{1}{\sum_{j=1}^{C} \frac{\exp\left(d_{in}/\beta\right)+1}{\exp\left(d_{jn}/\beta\right)+1}} \tag{17}$$

$$v_i = \frac{\sum_{n=1}^{N} u_{in} x_n}{\sum_{n=1}^{N} u_{in}} \tag{18}$$

It is obvious so far that the membership function of the nth entities provided by FCM, HCM and MEFCM algorithms depends upon the inverse of the square of the Euclidean distance $d_{in} = \|x_n - v_i\|^2$ which is a function of only x_n with no data or membership information of the clustering entity's neighbors are involved. Hence, the FCM, HCM and MEFCM algorithms miss important spatial local data and membership information. Thus additive noise can degrade x_n, v_i and d_{in}, thereby biasing the membership of a degraded entity to a false cluster.

5. HCM incorporating local membership KL divergence

In [18], an approach to incorporating local spatial membership information into HCM algorithm has been presented. By adding Kullback-Leibler (KL) divergence between the membership function of an entity and the locally-smoothed membership in the immediate spatial neighborhood, the modified objective function, called the local membership KL divergence-based FCM (LMKLFCM), is given by [18–22].

$$J_{LMKLFCM} = J_{HCM} + \gamma \left(\sum_{i=1}^{C} \sum_{n=1}^{N} u_{in} \log \left(\frac{u_{in}}{\pi_{in}} \right) + \sum_{i=1}^{C} \sum_{n=1}^{N} \overline{u}_{in} \log \left(\frac{\overline{u}_{in}}{\overline{\pi}_{in}} \right) \right) \quad (19)$$

where γ is a weighting parameter experimentally selected to control the fuzziness induced by the second term in (19), $\overline{u}_{in} = 1 - u_{in}$ is the complement of the membership function u_{in}, π_{in} and $\overline{\pi}_{in}$ are the spatial local or moving averages of membership u_{in} and the complement membership $\overline{u}_{in}$, functions respectively. These local membership and membership complement averages are computed by [18–22].

$$\pi_{in} = \frac{1}{N_K} \sum_{k \in N_n; k \neq n} u_{ik} \quad (20)$$

$$\overline{\pi}_{in} = \frac{1}{N_K} \sum_{k \in N_n; k \neq n} (1 - u_{ik}) = 1 - \pi_{in} \quad (21)$$

where N_n is a set of entities/pixels falling in a square window around the nth pixel and N_K is its cardinality. It is obvious that all entities in the window are weighted equally by $w_{pq}^{(u)} = 1/N_K$. Other windows can be used such as Gaussian one provided that the weight of the window-center is 0 and the rest weights are summed to unity. The first term in (19) provides hard-cluster labeling. It pushes the membership function toward 0 or 1. The KL membership and membership-complement divergences, in addition to providing fuzzification approach to HCM clustering, measure the proximity between the membership of a pixel in a certain cluster and the local average of the membership over the immediate spatial neighborhood pixels in this cluster. So, they push the membership function to the locally smoothed membership function π_{in}. Therefore, this can smooth out additive noise and bias the solution to piecewise homogenous labels which leads to a segmented image with piecewise homogenous regions.

The minimization of the objective function $J_{LMKLFCM}$ in (19) yields u_{in} and v_i to be given, respectively, by [18].

$$u_{in} = \frac{1}{\sum_{j=1}^{C} \left(\frac{\pi_{jn} \left((1-\pi_{in}) \exp(d_{in}/\gamma) + \pi_{in} \right)}{(1-\pi_{jn}) \exp(d_{jn}/\gamma) + \pi_{jn}} \right)} \pi_{in} = \delta_{in} \pi_{in} \quad (22)$$

$$v_i = \frac{\sum_{n=1}^{N} u_{in} x_n}{\sum_{n=1}^{N} u_{in}} \quad (23)$$

It is obvious from (22) that u_{in} is proportional to π_{in} and the proportional parameter δ_{in} is inversely proportional to the entity's distance d_{in} and the maximum δ_{kn} occurs when $d_{kn} = 0$.

It is clear that if $\gamma \to \infty$, $u_{in} = \pi_{in} / \sum_{j=1}^{C} \pi_{jn}$. Therefore, the resultant membership is independent of the data to be clustered but dependent on the initial value of the membership matrix U^0 and on the smoothing fashion. If u_{in}^0 is generated from a random process greater than zero, then u_{in}^t versus the number of iteration t converges, because of recursive averaging and normalizing, to a normal distribution variable with mean equal to $\frac{1}{C} = E\{u_{in}^t\} = E\{\pi_{in}\} / \sum_{j=1}^{C} E\{\pi_{jn}\}$ which, in this case, means too much fuzzy membership function. This has been proved experimentally by using a synthetic image of 4 clusters and $\gamma = 10^{10}$. Finally, as shown by (23), the computation of the cluster-center v_i is still independent of the local original data.

6. HCM incorporating local data and membership KL divergence

To incorporate local spatial data into the LMKLFCM objective function in (19), the following objective function has been proposed in [18].

$$J_{LDMKLFCM} = \sum_{i=1}^{C} \sum_{n=1}^{N} u_{in} \left(d_{in} + \alpha \bar{d}_{in} \right) + \gamma \left(\sum_{i=1}^{C} \sum_{n=1}^{N} u_{in} \log \left(\frac{u_{in}}{\pi_{in}} \right) + \sum_{i=1}^{C} \sum_{n=1}^{N} \bar{u}_{in} \log \left(\frac{\bar{u}_{in}}{\bar{\pi}_{in}} \right) \right) \tag{24}$$

Therefore, similar to (22) and (23), the membership function u_{in} and the cluster-center v_i are, respectively, given by [18].

$$u_{in} = \frac{1}{\sum_{j=1}^{C} \left(\frac{\pi_{jn} \left((1-\pi_{in}) \exp\left(\left(d_{in} + \alpha \bar{d}_{in} \right) / \gamma \right) + \pi_{in} \right)}{\left((1-\pi_{jn}) \exp\left(\left(d_{in} + \alpha \bar{d}_{in} \right) / \gamma \right) + \pi_{jn} \right)} \right)} \pi_{in} \tag{25}$$

$$v_i = \frac{\sum_{n=1}^{N} u_{in} (x_n + \alpha \bar{x}_n)}{(1+\alpha) \sum_{n=1}^{N} u_{in}} \tag{26}$$

It is obvious that the LDMKLFCM algorithm in (24)–(26) provides a membership that depends upon the local spatial data and membership information while the cluster center is dependent upon the locally-smoothed data. Thus the algorithm has twofold approach to handle additive noise.

7. Simulation results

This simulation aims at examining the performance of the conventional FCM, the membership entropy-based FCM (MEFCM), the spatial distance weighted FCM (SFCM), the local membership KL divergence-based FCM (LMKLFCM) and the local data and membership KL

divergence-based FCM (LDMKLFCM) algorithms. It is to be noticed that all the algorithms can be implemented almost similar to the pseudo code in **Table 1** by replacing the steps 3 and 4 by the corresponding computation of the membership function and cluster centers of each algorithm.

7.1. Clustering validity

To measure the performance of the fuzzy clustering algorithms, several quantitative measures or indices have been adopted in [23, 25] and references therein. Few of these measures are the partition coefficient V_{PC} and the partition entropy V_{PE} index of Bezdek and Xie-Beni (XB index) V_{XB}, given respectively by

$$V_{PC} = \frac{1}{N}\sum\nolimits_{n=1}^{N}\sum\nolimits_{i=1}^{C} u_{in} \tag{27}$$

$$V_{PE} = -\frac{1}{N}\sum\nolimits_{n=1}^{N}\sum\nolimits_{i=1}^{C} u_{in}\log(u_{in}) \tag{28}$$

The closer of the V_{PC} to 1, the better the performance since the minimization is constrained by $\sum_{i=1}^{C} u_{in} = 1$. The closer the V_{PE} to 0, the better the performance since this means the less fuzziness of the membership and thus clusters are well-separated.

In synthetic images, in addition to the above clustering validity measures, several clustering validity and performance measures have also been used such as the accuracy, sensitivity and specificity given respectively by

$$Acc. = (TP + TN)/(TP + TN + FP + FN) \tag{29}$$

$$Sen. = TP/(TP + TN) \tag{30}$$

$$Spe. = TN/(TN + FN) \tag{31}$$

where T, F, P, and N are mean true, false, positive, and negative, respectively. The TP, FP, TN, and FN are computed as follows. While generating the synthetic image, the ground truth labels are formulated as the logical matrix given by [23].

$$L_{in} = \begin{cases} 1; if\ x_n \in i \\ 0; otherwise \end{cases} ;\ i = 1, 2, \ldots, C, n = 1, 2, .., N. \tag{32}$$

where x_n is the noise-free pixel in the synthetic image and 1 and 0 represent True and False, respectively. After the segmentation is done, the estimated labels are also formulated as logical matrices generated by [20].

$$\widehat{L}_{kn} = \begin{cases} 1; k = \arg\max_i(u_{in}) \\ 0; otherwise \end{cases} ;\ i = 1, 2, \ldots, C, n = 1, 2, .., N. \tag{33}$$

Finally, the TP, TN, FP, and FN are given by [20].

$$TP = \sum_{i=1}^{C} \sum_{n=1}^{N} \widehat{L}_{in} L_{in}; \qquad TN = \sum_{i=1}^{C} \sum_{n=1}^{N} \overline{\widehat{L}}_{in} \overline{L}_{in}$$

$$FP = \sum_{i=1}^{C} \sum_{n=1}^{N} \widehat{L}_{in} \overline{L}_{in}; \qquad FN = \sum_{i=1}^{C} \sum_{n=1}^{N} \overline{\widehat{L}}_{in} L_{in}$$

(34)

where "＿" means the logical complement.

7.2. Artificial image

In this simulation, the artificial or synthetic noise-free image shown in **Figure 1(a)** is degraded by adding zero-mean white Gaussian noise (WGN) with different variances. The noisy image
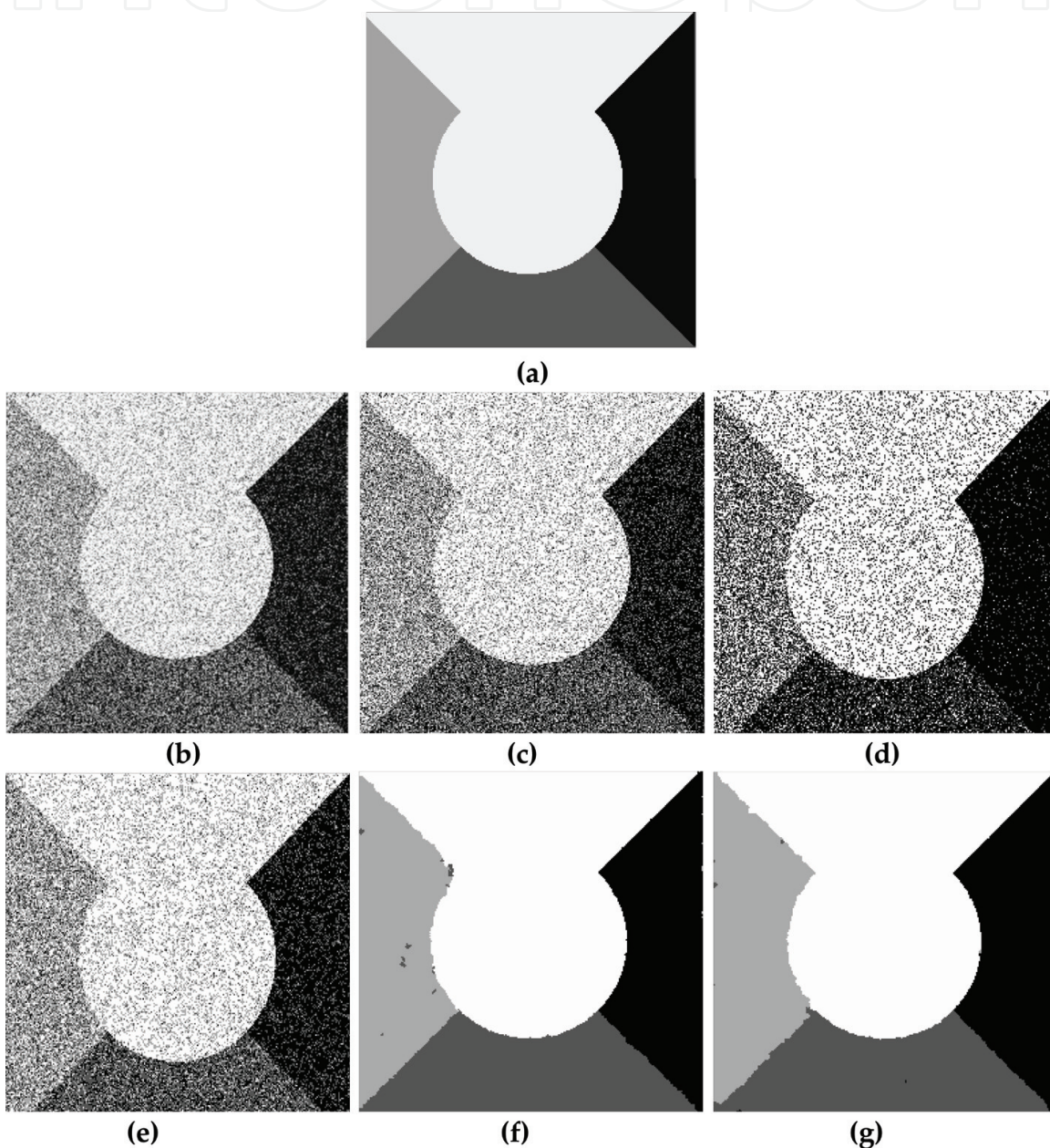

Figure 1. Clustering of the synthetic image: (a), noise free-image; (b), the noise-free image plus zero-mean and 0.08 variance WGN; (c) FCM; (d), MEFCM; (e), SFCM; (f), LMKLFCM; (g), LDMKLFCM. It is evident that the clustered images in (f) and (g) have lesser number of misclassified pixels which means that noisy pixels are rightly clustered. Clustering validation measures are summarized in **Table 2**.

shown in **Figure 1(b)** is for 0.08 noise variance. We have studied the performance of the five algorithms, namely, the standard FCM, the membership entropy-based FCM (MEFCM), the spatial distance weighted FCM (SFCM), the local membership KLFCM (LMKLFCM) and the local data and membership KLFCM (LDMKLFCM) algorithms in segmenting these noisy images with $m = 2$ and $C = 4$. The parameters for the algorithms have been elected via simulation as $\beta = 1000$ for MEFCM; $\lambda = 0.5$ for SFCM; $\gamma = 1000$ for LMKLFCM; and $\gamma = 1000$ and $\alpha = 0.5$ for LDMKLFCM. For the computation of the locally smoothed data $\bar{x}_n$, a neighboring window of size 3x3 has been used. Also, the same spatial window has been used for the computation of the locally-smoothed membership function π_{in}. The initial values of the membership functions U and the cluster-centers V are generated from a uniformly distributed random process with means 0.5 and equal to the image mean, respectively. We have collected results from 25 Monte Carlo runs of each algorithm. In each run, the initial values of U and V of the FCM are new random samples while the ones of the rest algorithms are generated by executing few number of iterations of the FCM algorithm. Simulation results, not included for space limitation, have shown that the algorithms provide further improvement with these initial values generated by the FCM algorithm than those randomly generated. Also, in each run, a new random sample of WGN is used in generating the noisy images. **Figure 1(c–g)** show the clustered images generated by the five algorithms in the case of 0.08 noise variance. These clustered images show that the LMKLFCM and the LDMKLEFCM algorithms provide the ones with lesser noise which means lesser number of misclassified pixels. Moreover, the LDMKLFCM algorithm offers the superior clustered image. **Table 2** summarizes the averages and standard deviations ($\mu \pm \sigma$) of the performance measures. The LMKLFCM and LDMKLFCM show the maximum V_{PC} and the minimum V_{PE}. The averages of the accuracy, sensitivity and the specificity performance measures of the five algorithms have been studied

| Algorithm	Images	V_{PC}	V_{PE}
FCM	Synthetic	0.8105 ±0.0007	0.3517 ± 0.0012
	Simulated MR	0.7921 ±0.0011	0.3986 ± 0.0020
	Real MR	0.8930 ±0.0140	0.1998 ± 0.0240
	Lena	0.8286 ±0.0004	0.2824 ± 0.0006
SFCM	Synthetic	0.8370 ± 0.0010	0.3017 ± 0.0017
	Simulated MR	0.8674 ± 0.0009	0.2409 ± 0.0014
	Real MR	0.9204 ± 0.0006	0.1440 ± 0.0012
	Lena	0.8936 ± 0.0006	0.1786 ± 0.0009
MEFCM	Synthetic	0.8616 ± 0.0012	0.2271 ± 0.0019
	Simulated MR	0.8873 ± 0.0012	0.1841 ± 0.0018
	Real MR	0.9602 ± 0.0113	0.0650 ± 0.0183
	Lena	0.9268 ± 0.0004	0.1198 ± 0.0007
LMKLFCM	Synthetic	0.9853 ± 0.0011	0.0270 ± 0.0028
	Simulated MR	0.8958 ± 0.0088	0.1721 ± 0.0146
	Real MR	0.9625 ± 0.0087	0.0441 ± 0.0128
	Lena	0.9609 ± 0.0012	0.0643 ± 0.0020
LDMKLFCM	Synthetic	0.9874 ± 0.0011	0.0227 ± 0.0022
	Simulated MR	0.9234 ± 0.0030	0.1258 ± 0.0049
	Real MR	0.9519 ± 0.0016	0.0604 ± 0.0025
	Lena	0.9730 ± 0.0026	0.0446 ± 0.0026

Table 2. Clustering validation measures for synthetic and real-world images.
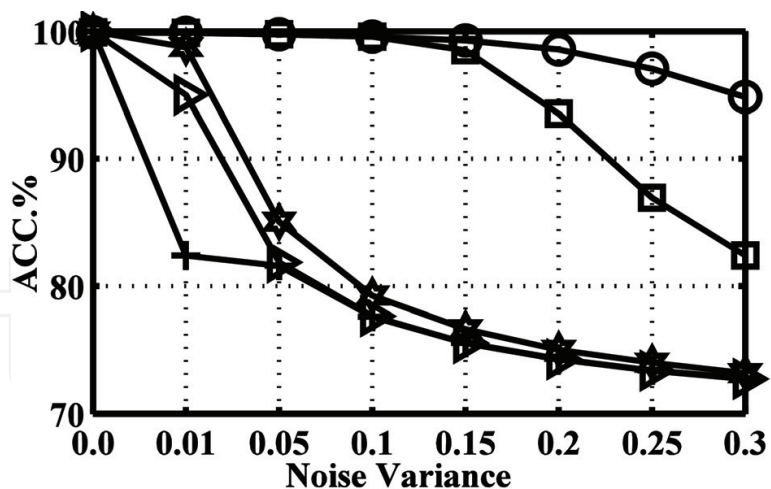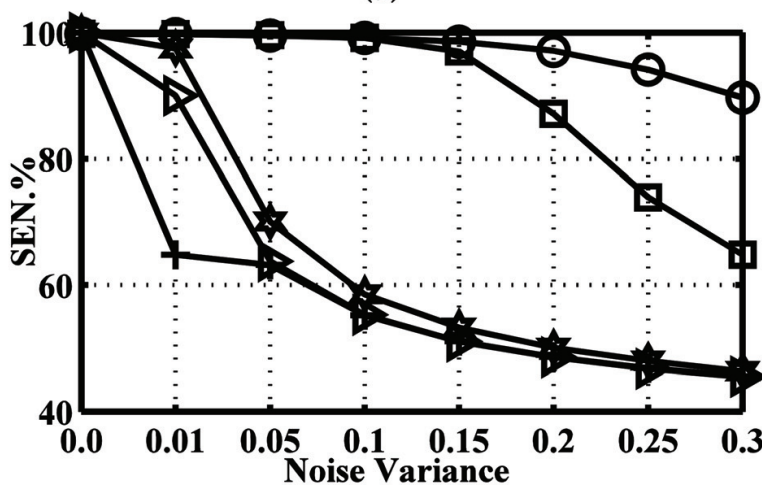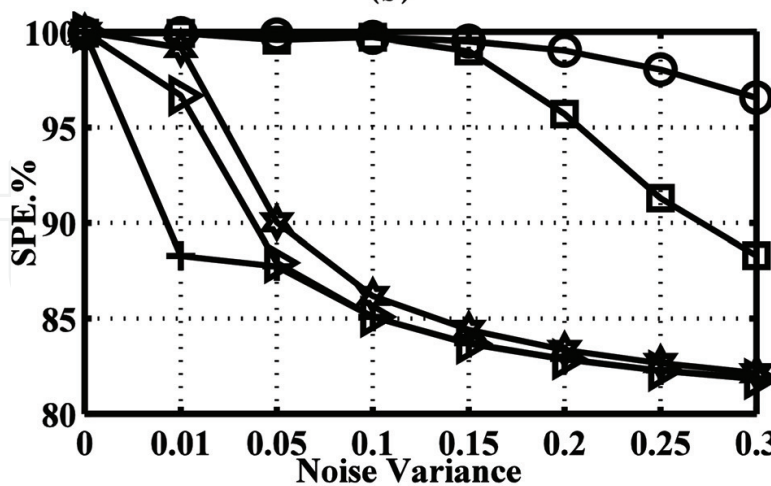
Figure 2. The average versus noise variance of accuracy, (a); sensitivity, (b); and specificity, (c); ▷, FCM; +, MEFCM; SFCM; LMKLFCM; LDMKLFCM. The proposed LMKLFCM and LDMKLFCM algorithms provide the superior performance among the five algorithms. The LDMKLFCM algorithm shows more noise-robust capability.
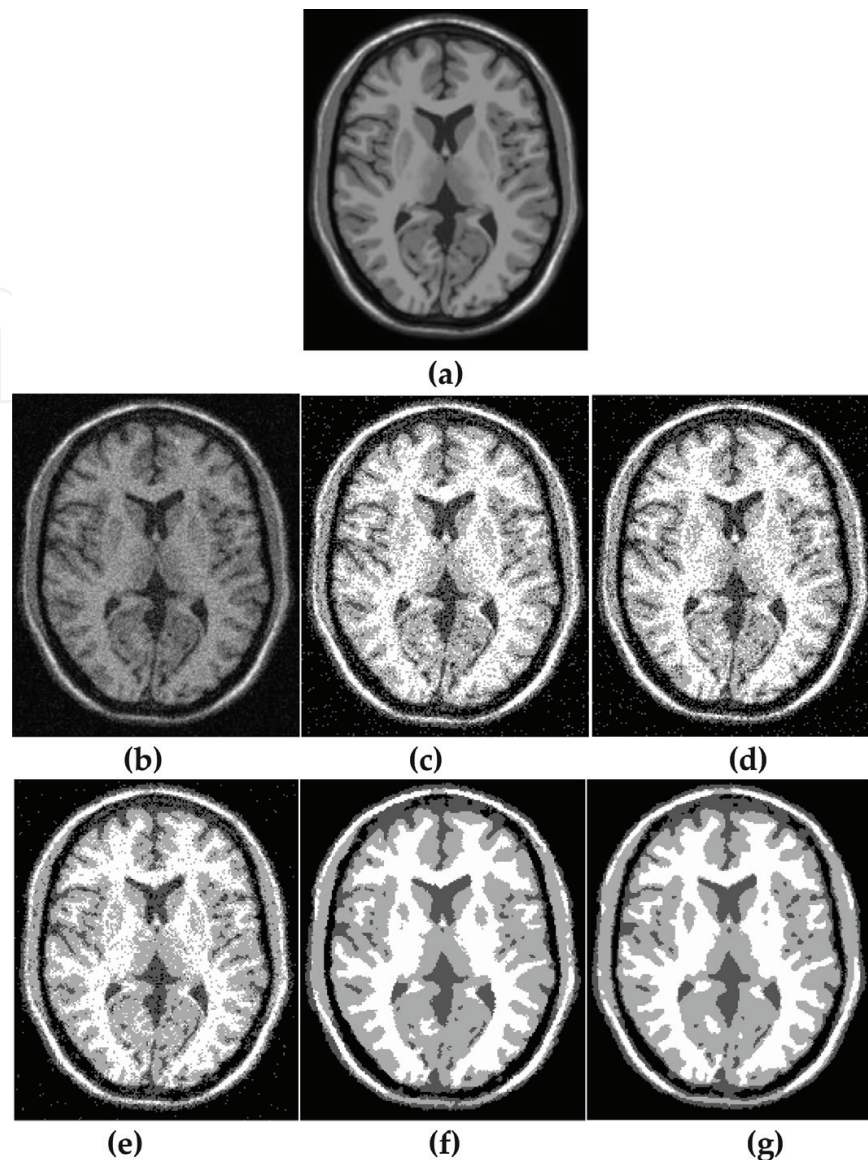
Figure 3. Clustering of simulated MRI: (a), noise-free MRI; (b), the MRI in (a) plus zero-mean WGN with 0.005 variance. Segmented images by: (c), FCM; (d), MEFCM; (e), SFCM; (f), LMKLFCM (g), LDMKLFCM. Obviously, the segmented images in (f) and (g) provided by the LMKLFCM and the LDMRKLCM algorithms, respectively, have lesser noise which means that the noisy pixels are correctly clustered. The clustering validation measures summarized in **Table 2** show that the LMRKlCM; and LDMKLFCM provide the maximum V_{PC} and the minimum V_{PE}.

against noise variance. **Figure 2** shows these measures versus noise variance. It is clear that both the LMKLFCM and the LDMKLFCM algorithms provide the superior performance among the five algorithms and the LDMKLFCM algorithm shows more noise-robustness.

7.3. Magnetic resonance image (MRI)

A simulated MRI of [26], illustrated by **Figure 3(a)**, has been used as a noise-free image. It has been degraded by adding white Gaussian noise (WGN) with zero-mean and 0.005 variance to
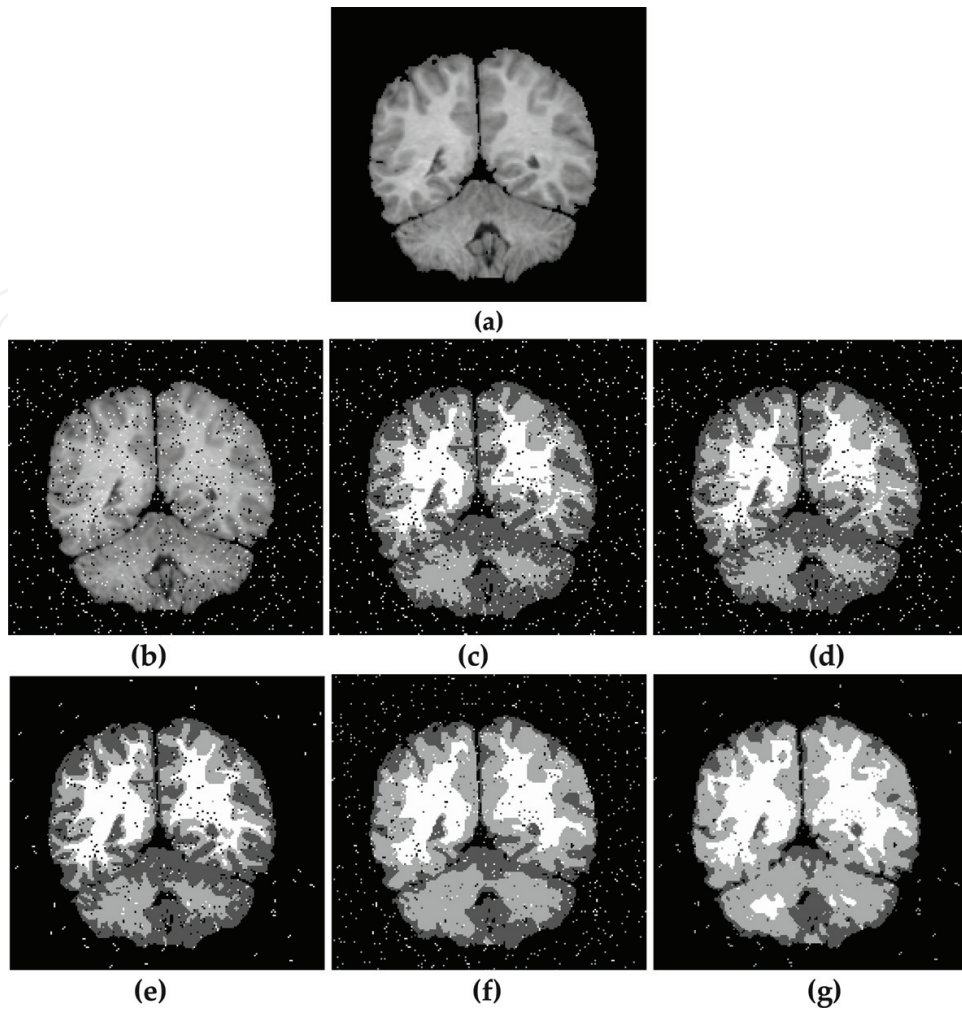
Figure 4. Clustering of real MRI example: (a), noise-free real MRI; (b), the image in (a) plus salt&pepper with 0.05 variance. Segmented images by: (c), FCM; (d), MEFCM; (e), SFCM; (f), LMKLFCM (g), LDMKLFCM. Clearly, the segmented images in (f) and (g) generated by the LMKLFCM and the LDMRKLCM algorithms, respectively, have lesser noise. The clustering validation coefficients summarized in **Table 2** show that the LMRKlCM; and LDMKLFCM provide the maximum V_{PC} and the minimum V_{PE}.

generate the noisy MRI illustrated by **Figure 3(b)**. This noisy MRI image has been clustered by the five algorithms. The parameters for all algorithms have been taken similar to the ones of the synthetic image simulation except, for the MEFCM algorithm, $\beta = 200$ and, for both LMKLFCM and LDMKLFCM algorithms, $\gamma = 1000$. We have also executed 25 runs of each algorithm. The initial values of u_{in} and v_i have been generated and adjusted as explained in the synthetic image simulation. **Figure 3(c–g)** shows the resulting clustered images provided by the five algorithms in a certain run. **Table 2** shows the averages and standard deviations ($\mu \pm \sigma$) of the performance measures V_{PC} and V_{PE} of the five algorithms. It obvious that the LMKLFCM and LDMKLFCM provide the segmented images with lesser noise or lesser number of misclassified pixels, the maximum V_{PC} and the minimum V_{PE}.

A real MRI from [27], shown in **Figure 4(a)**, has been considered as a noise-free image. To generate the noisy MRI shown in **Figure 4(b)**, salt & pepper noise with 0.050 variance have

been added. The noisy MRI has been clustered by the FCM, SFCM, MEFCM, LMKLCM and the LDMKLFCM algorithms. The parameters for all algorithms have been taken similar to the ones of the synthetic image simulation except, for the MEFCM algorithm, $\beta = 300$ and, for both the LMKLFCM and LDMKLFCM algorithms, $\gamma = 800$. We have also obtained the results of 25 runs of each algorithm. The initial values of u_{in} and v_i have been generated and adjusted as

Figure 5. Segmentation of Lena image: (a), noise-free image; (b), the image in (a) plus WGN noise with zero-mean and 0.05 variance. It is obvious that the images in (f) and (g) have lesser number of misclassified pixels. The clustering validation coefficients summarized in **Table 2** which shows that the LMKLFCM and the LDMKLFCM algorithms provide the superior V_{PC} and V_{PE}.

mentioned in the synthetic image simulation. **Figure 4(c–g)** show the segmented images provided by the five algorithms in a certain run while **Table 2** summarizes the averages and standard deviations ($\mu \pm \sigma$) of the performance measures. It is obvious that the proposed LMKLFCM and LDMKLFCM algorithms provide the segmented images with lesser noise or lesser number of misclassified pixels, the maximum V_{PC} and the minimum V_{PE}.

7.4. Lena image

A popular Lena image shown in **Figure 5(a)** has been considered as a noise-free image example. The noisy Lena image shown in **Figure 5(b)** has been generated by adding WGN noise with zero-mean and 0.01 variance. The parameters of the five algorithms have been adjusted to the values similar to the ones used in the previous simulations except $C = 2$; $\beta = 1000$ for the MEFCM algorithm; $\gamma = 2000$ for the LMREFCM and $\gamma = 2000$ and $\alpha = 0.5$ for the LDMREFCM algorithms. We have also executed 25 Mont Carlo Runs of each algorithm as explained above. **Figure 5(c–g)** shows the resulting segmented images obtained by the five algorithms. Visually investigation of the segmented images shows that the LMKLFCM and LDMKLFCM algorithms provide the images with lesser number of misclassified pixels. **Table 2** shows the average and standard deviation ($\mu \pm \sigma$) of the performance measures of the five algorithms. It is also clear that the two algorithms provide the maximum V_{PC} and the minimum V_{PE}.

8. Conclusions

The hard C-means algorithm has been fuzzified by incorporating into the objective function spatial local information through two KL membership divergences. The first KL membership divergence measures the information proximity between the membership of each pixel and its local membership average in the pixel neighborhood. The second one measures the information proximity between the complement membership and its local membership average in the pixel neighborhood. For regularization, the local data information has been incorporated by an additional new weighted hard C-means function in which the noisy-image is replaced by a noise-reduced one. Such incorporation of both local data and local membership information facilitates biasing the algorithm to classify each pixel in correlation with its immediate neighboring pixels. Results of segmentation of synthetic, simulated medical and real-world images have shown that the proposed local membership KL divergence-based FCM (LMKLFCM) and the local data and membership KL divergence-based entropy FCM (LDMKLFCM) algorithms outperform several widely used FCM related algorithms. Moreover, the average runtimes of all algorithms have been measured via simulation. In all runs, all algorithms start from the same randomly generated initial conditions, as mentioned in the simulation section, and stopped at the same fixed point. The LDMKLFCM, LMKLFCM, standard FCM, MEFCM, and SFCM algorithms have provided average runtime of 1.5, 1.75, 1, 0.9 and 1 sec respectively. The simulation results have been done using Matlab R2013b under windows on a processor of Intel (R) core (TM) i3, CPU M370 2.4 GHZ, 4 GB RAM.

Acknowledgements

Conflict of interest

No potential conflicts of interest to report.

Author details

Reda R. Gharieb[1,2]*

*Address all correspondence to: rrgharieb@gmail.com

1 Faculty of Engineering, Assiut University, Assiut, Egypt

2 Higher Institute of Engineering, Thebes Academy for Sciences, Cairo, Egypt

References

[1] Pal NR, Pal SK. A review on image segmentation techniques. Pattern Recognition. 1993; **26**(9):1277-1294

[2] Ravindraiah R, Tejaswini KA. Survey of image segmentation algorithms based on expectation-maximization. IOSR Journal of VLSI and Signal Processing (IOSR-JVSP). 2013;**2**:01-07

[3] Pham DL, Xu C, Prince JL. Current methods in medical image segmentation. Annual Review of Biomedical Engineering. 2000;**2**:315-337

[4] Jain AK. Data clustering: 50 years beyond K-means. Pattern Recognition Letters. 2010;**3**: 651-666

[5] MacQueen J. Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1967;**14**:281-297

[6] Alsabti K, Ranka S, Singh V: An efficient k-means clustering algorithm. Electrical Engineering and Computer Science. Paper 43

[7] Bezdek JC. Pattern Recognition with Objective Fuzzy Algorithms. New York: Plenum Press; 1981

[8] Zadeh LA. Fuzzy sets. Information and Control. 1965;**8**:338-353

[9] Chuang KS, Tzeng HL, Chen S, Wu J, Chen TJ. Fuzzy c-means clustering with spatial information for image segmentation. Computerized Medical Imaging and Graphics. 2005;**30**:9-15

[10] Miyamoto S, Ichihashi H, Honda K. Algorithms for Fuzzy Clustering. Heidelberg: Springer; 2008

[11] Ahmed MN, Ymany S, Mohamed N, Farag A, Moriarty T. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. IEEE Transactions on Medical Imaging. 2002;**21**:193-199

[12] Krinidis CSV. A robust fuzzy local information and C-means clustering algorithm. IEEE Transactions on Image Processing. 2010;(5):1328-1337

[13] Guo Y, Liu K, Wu Q, Hong Q, Zhang H. A new spatial fuzzy c-means for spatial clustering. Wseas Transactions on Computer. 2015;**14**:369-381

[14] Cai W, Chen S, Zhang D. Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation. Pattern Recognition. 2005;**40**:825-838

[15] Honda K, Ichihashi H. A new approach to fuzzification of memberships in cluster analysis. Modeling Decisions for Artificial Intelligence. 2005:97-124

[16] Yao J, Dash M, Tan ST, Liu H. Entropy-based fuzzy clustering and fuzzy modeling. Fuzzy Sets and Systems. 2000;**113**:381-388

[17] Yasuda M. Fuzzy c-Means Clustering, Entropy Maximization, and Deterministic and Simulated Annealing. InTech Book of Simulated Annealing-Advances, Applications and Hybridizations. 2012. DOI: 10.5772/48659

[18] Gharieb RR, Gendy G, Abdelfattah A. C-means clustering fuzzified by two membership relative entropy functions approach incorporating local data information for noisy image segmentation. Signal, Image and Video Processing. 2017;**11**:541-548. https://doi.org/10.1007/s11760-016-0992-4

[19] Gharieb RR, Gendy G, Selim H. A hard C-means clustering algorithm incorporating membership KL divergence and local data information for noisy image segmentation. International Journal of Pattern Recognition and Artificial Intelligence. 2017:1-12. https://doi.org/10.1142/S021800141850012X

[20] Gharieb RR, Gendy G, Abdelfattah A, Selim H. Adaptive local data and membership based KL divergence incorporating C-means algorithm for fuzzy image segmentation. Applied Soft Computing. 2017. https://doi.org/10.1016/j.asoc.2017.05.055

[21] Gharieb RR. Data Science- Scientific and Statistical Computing. Germany: Noor Publishing; 2017

[22] Gharieb RR, Gendy G. Fuzzy C-means with a local membership KL distance for medical image segmentation. In: Proceedings of the IEEE International Conference on Biomedical Engineering Conference (CIBEC 14); 11–13 December 2014; Cairo; IEEE; p. 47-50

[23] Pal NR, Bezdek JC. On cluster validity for the fuzzy c-means model. IEEE Transactions on Fuzzy Systems. 1995;**3**:370-379

[24] Gharieb RR. Gendy fuzzy C-means with local membership based weighted pixel distance and KL divergence for image segmentation. Journal of Pattern Recognition Research. 2015;**1**:53-60

[25] Wang W, Zhang Y. On fuzzy cluster validity indices. Fuzzy Sets and Systems. 2007;**158**: 2095-2117

[26] Online simulated brain web. Available at: http://brainweb.bic.mni.mcgill.ca/brainweb/

[27] Internet brain segmentation repository (ibsr). Available at: https://www.nitrc.org/projects/ibsr