ORIGINAL PAPER

Noise-resistant fuzzy clustering algorithm

S. Askari[1] ●

Abstract

The main objective of Fuzzy C-means (FCM) algorithm is to group data into some clusters based on their similarities and dissimilarities. However, noise and outliers affect the performance of the algorithm that results in misplaced cluster centers. Although several corrections are made in the algorithm to tackle this problem but the algorithm is not improved effectively and still suffers from the same problem. Noise-resistant FCM (nrFCM) algorithm is proposed in this work to improve the performance of the FCM algorithm when dealing with noise and outliers. The nrFCM algorithm eliminates the effects of noise and outliers on the cluster centers by introducing a function of distance instead of the distance itself into the objective function of the FCM algorithm. It is shown that the nrFCM algorithm is significantly more accurate than the FCM algorithm and noise and outliers cannot impair its accuracy. However, its runtime is higher than that of the FCM algorithm because of nonlinear update equation for cluster centers.

Keywords Fuzzy C-means (FCM) · Fuzzy clustering · Outlier · Noise · Possibilistic C-means (PCM) · Possibilistic fuzzy C-means (PFCM)

1 Introduction

Granular computing is a framework for data and information processing that includes several methods of soft computing, machine learning, and artificial intelligence (Dubois and Prade 2016; Wang et al. 2017; Livi and Sadeghian 2016; Ciucci 2016). Granular computing is applied to various problems such as fuzzy systems (Apolloni et al. 2016; Liu and Zhang 2018), classification (Amezcua and Melin 2019; Liu and Cocea 2017; Aydav and Minz 2020; Antonelli et al. 2016; Liu and Cocea 2019), traffic scene recognition (Wu et al. 2019), situation awareness (Loia et al. 2016), and clustering (Chen et al. 2019b; Zhang et al. 2019a; Martino and Sessa 2020; Lingras et al. 2016; Peters and Weber 2016). Fuzzy clustering algorithms are designed to identify compact groups of data points within the data. In recent years, these algorithms have been applied to various research areas. For instance, a

color image can be segmented by fuzzy clustering (Ozdemir and Akarun 2011; Tolias and Panas 1998), where the objects are presented by clusters with different characteristics to preserve details of the image. It is shown that fuzzy clustering effectively determines spatial organization and compactness of pixel clusters (Beliakov et al. 2015). Type-II Fuzzy C-Means (FCM) algorithm is utilized to extract different features of scene images for a humanoid robot that shows the superior performance of this approach over other methods (Liu et al. 2014). A segmentation approach is presented based on fuzzy clustering and it is shown that the method is useable in object recognition, image retrieval, and human vision simulation (Makrogiannis et al. 2005). Fuzzy time series is employed for simulating behavior of imprecise data with a small number of samples and unclear trend (Li and Cheng 2010; Wong et al. 2010). Although some of the forecasting algorithms presented for such time series are based on intervals of the universe of discourse of the input variables (Chen and Jian 2017; Chen and Phuong 2017; Chen et al. 2012, 2013, 2019a; Chen and Chen 2015; Chen and Wang 2010), but it is shown that fuzzy clustering provides more accurate predictions for fuzzy time series (Zeng et al. 2019; Chen and Chang 2010; Chen and Tanuwijaya 2011a, b; Askari and Montazerin 2015; Cheng et al. 2008; Egrioglu et al. 2011; Aladag et al.

✉ S. Askari
  s_askari@aut.ac.ir; bas_salaraskari@yahoo.com

1 Mechanical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), 424 Hafez Avenue, Tehran 1591634311, Iran

2012; Chen and Chen 2014; Duru and Bulut 2014; Askari et al. 2015a, b). It is shown that oil reservoir with noisy data is efficiently classified by fuzzy clustering algorithms (Askari 2017a). Because of high complexity, network type problems such as gas and water distribution networks and electrical grids require fast forecasting systems to predict their behavioral. (Askari et al. 2016a, b). It is shown that fuzzy systems based on fuzzy clustering are suitable choices for such problems (Askari et al. 2020). Fuzzy clustering is also used for system identification (Montazerin et al. 2015) and structure identification of fuzzy systems (Tung and Quek 2004; Zhang et al. 2002; Askari 2017b).

The FCM algorithm is the basic fuzzy clustering technique (Bezdek et al. 1984; Hathaway and Bezdek 2001) and its convergence proof is presented in (Groll and Jakel 2005). So far, many developments are made on the algorithm to overcome its deficiencies. The algorithm is developed to handle very large or big data that cannot be loaded in the working memory of a computer (Havens et al. 2012). A multiple-kernel FCM algorithm is proposed for image segmentation by linear combination of various kernels (Chen et al. 2011). The algorithm fuses different pixel information captured by different kernels in the image segmentation (Chen et al. 2011). Applying the algorithm to synthetic and medical images proves its advantages. Degree of fuzziness is typically set to 2.0 in fuzzy clustering algorithms, however, changing this parameter may seriously affect results of clustering. Large value of this parameter increases fuzziness and overlap of the clusters but when it tends to 1.0, the clusters become crisp. A generalized form of the FCM algorithm is proposed and its robustness and convergence are analyzed (Zhu et al. 2009). The algorithm shows some advantages over the FCM algorithm when applied to noisy images (Zhu et al. 2009). Fuzzy sets and rough sets are combined to utilize advantages of both concepts in fuzzy clustering (Maji and Pal 2007). Qualitative and quantitative examinations of this hybrid algorithm demonstrate its effectiveness when it is applied to real-life data.

The main objective of the present work is to formulate a fuzzy clustering algorithm to cancel impacts of noise and outliers on the cluster centers. To this purpose, the distance in the objective function of the FCM algorithm is replaced with an exponential function. As a result, an exponential function of distance appears in the update equation of the cluster centers. Since noise and outliers are mostly far away from the clusters centers, their impacts on the clustering results become negligible. Effectiveness of this idea is proved by several experiments carried out on the noisy data.

Rest of the paper is organized as follows. Research works closely related to the present study are briefly discussed in Sect. 2. Several algorithms are discussed in Sect. 3 along with the proposed algorithm. Performances of these algorithms on noisy data are compared in Sect. 4 and summary of the paper and its findings are provided in Sect. 5.

2 Related works

Sensitivity to noise and outliers is one of the main drawbacks of the FCM algorithm. The problem is addressed in many research works. The Credibilistic fuzzy C-means (CFCM) algorithm takes the credibility of data points to the whole data into account to reduce the sensitivity of the FCM algorithm to outliers (Chintalapudi and Kam 1998). Performance of this algorithm proves its superiority over the FCM algorithm. A noise-tolerant FCM algorithm is proposed for data stream regression (Song et al. 2019). It is shown that this algorithm provides more accurate regressions as compared to several well-known methods. The Density Oriented Fuzzy C-Means (DOFCM) algorithm is presented to identify outliers in the data (Kaur and Gosain 2010). Location of the cluster centers calculated by the DOFCM algorithm is not affected by noise and outliers. A robust fuzzy clustering algorithm is proposed based on membership linking and filtering for segmentation of noise images (Lei et al. 2018; Wang et al. 2020). These algorithms incorporate pixels labels into the clustering algorithm to apply to the supervised process of image segmentation. Performance of these algorithms is compared to that of several methods and it is shown that they are significantly more accurate than the other algorithms. Type-2 FCM algorithm is proposed to capture uncertainty and fuzziness of the data more efficiently (Gosain and Dahiya 2016). Since data points with higher membership grades determine the cluster centers calculated by this algorithm, it is supposed to be capable of removing noise and outliers impacts on the clustering results. The algorithm is further extended by replacing the distance in the objective function of the original algorithm with Gaussian kernels. This modification significantly improves performance of the algorithm when dealing with noise and outliers.

Possibilistic C-Means (PCM) algorithm is proposed to handle noise and outliers (Krishnapuram and Keller 1993, 1996; Koutroumbas et al. 2018; Anderson et al. 2010; Zhang et al. 2017a, b; Zhang and Leung 2004; Pal et al. 2005). Although it is applied to several problems such as boundary detection and surface approximation (Krisnapuram et al. 1995a, b), image segmentation (Kalist et al. 2015; Zhang et al. 2011; Xie et al. 2007), big data clustering (Zhang et al. 2019a, b), sensor big data (Zhang and Chen 2014), cloud computing (Srinivasan and Palanisamy

2015), system identification (Ahmed et al. 2015; Bouzbida et al. 2017), and outlier detection (Filippone et al. 2010), but the algorithm severely suffers from sensitivity to initialization and coincident clusters (Koutroumbas et al. 2018; Askari et al. 2017a, b.). To overcome these problems, the FCM and PCM algorithms are combined to make the hybrid Possibilistic Fuzzy C-Means (PFCM) algorithm to take advantages of both algorithms and avoid their deficiencies (Pal et al. 2005).

However, the PFCM algorithm does not provide the expected results and the cluster centers calculated by this algorithm are highly influenced by noise and outliers. For example, consider the data shown in Fig. 1 that are used in (Pal et al. 2005) (where the PFCM algorithm is introduced for the first time) to show the effectiveness of the PFCM algorithm in handling outliers. The data point $\begin{bmatrix} 0 & 0 \end{bmatrix}^T$ is inlier and $\begin{bmatrix} 0 & 10 \end{bmatrix}^T$ is outlier. These data are clustered using the PCM, FCM, and PFCM algorithms and the cluster centers computed by these algorithms are indicated by $\diamond$, $+$, and $\times$, respectively. The PFCM algorithm obviously fails to find the accurate cluster centers because of these two points added to the original data that contains two diamond-shaped clusters. Similarly, the PCM and FCM algorithms calculate misplaced cluster centers. The cluster centers indicated by * are calculated by the noise-resistant FCM (nrFCM) algorithm presented in this paper that is discussed later. The nrFCM algorithm initializes by the FCM algorithm such that the initial cluster centers are calculated by the FCM algorithm and then the nrFCM algorithm updates these prototypes to find the final cluster centers.
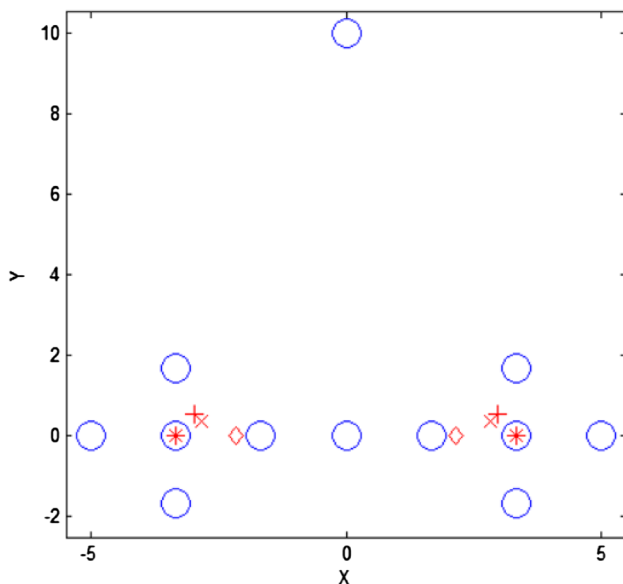

Fig. 1 Cluster centers of the data with two clusters calculated by the PCM, FCM, PFCM, and nrFCM algorithms

3 Theoretical background

Three algorithms are discussed prior to nrFCM formulation including the FCM algorithm, Type-2 Fuzzy C-Means (T2FCM), and Kernelized Type-2 Fuzzy C-Means (KT2FCM)and. These algorithms are later used to evaluate the performance of the nrFCM algorithm.

3.1 Fuzzy C-means (FCM) algorithm

Objective function of the FCM algorithm is as follows (Pal et al. 2005) where c is the number of clusters, n is the number of data points, $\vec{x}_j$ is the jth data point, and $\vec{v}_i$ is the ith cluster center. Moreover, m is degree of fuzziness and u_{ij} is membership grade of $\vec{x}_j$ in $\vec{v}_i$.

$$J = \sum_{j=1}^{n} \sum_{i=1}^{c} u_{ij}^m d_{ij}^2, \sum_{k=1}^{c} u_{kj} = 1 \tag{1}$$

where $d_{ij}^2 = (\vec{x}_j - \vec{v}_i)^T A (\vec{x}_j - \vec{v}_i)$ is distance of $\vec{x}_j$ from $\vec{v}_i$ and $A_{r \times r}$ is a norm matrix where r is dimension of the data. Type of the norm matrix determines shape of the clusters. Identity norm matrix $I_{r \times r}$ generates spherical clusters and the following covariance norm matrix provides elliptic clusters.

$$A = \left(\frac{1}{n} \sum_{j=1}^{n} (\vec{x}_j - \bar{v})(\vec{x}_j - \bar{v})^T \right)^{-1}, \bar{v} = \frac{1}{n} \sum_{j=1}^{n} \vec{x}_j \tag{2}$$

The membership grades and cluster centers are calculated by minimizing (1) that results in the following update equations for the FCM algorithm (Pal et al. 2005).

$$u_{ij} = \left[\sum_{k=1}^{c} \left(\frac{d_{ij}^2}{d_{kj}^2} \right)^{\frac{1}{m-1}} \right]^{-1}, \vec{v}_i = \frac{\sum_{j=1}^{n} u_{ij}^m \vec{x}_j}{\sum_{j=1}^{n} u_{ij}^m} \tag{3}$$

3.2 Type-2 fuzzy C-means (T2FCM)

The T2FCM algorithm is based on Type-2 fuzzy sets. This algorithm calculates cluster centers differently as compared to the FCM algorithm such that data points with higher membership grades are more determinative of the final position of the cluster centers. Since membership grades of noise and outliers are less than those of actual data points, the T2FCM algorithm is supposed to show good performance on noisy data (Gosain and Dahiya 2016). Type-1 membership grades of this algorithm are calculated from (3) as follows.

$$u_{ij} = \left[\sum_{k=1}^{c} \left(\frac{d_{ij}^2}{d_{kj}^2} \right)^{\frac{1}{m-1}} \right]^{-1} \tag{4}$$

The Type-2 membership grades μ_{ij} are then calculated from the Type-1 membership grades by the following equation (Gosain and Dahiya 2016).

$$\mu_{ij} = u_{ij} - \frac{1 - u_{ij}}{2} \tag{5}$$

Finally, cluster centers are calculated as follows (Gosain and Dahiya 2016).

$$\vec{v}_i = \frac{\sum_{j=1}^n \mu_{ij}^m \vec{x}_j}{\sum_{j=1}^n \mu_{ij}^m} \tag{6}$$

3.3 Kernelized type-2 fuzzy C-means (KT2FCM)

The KT2FCM algorithm is proposed by further generalization of the T2FCM algorithm by replacing the distance in (1) with hyper tangent kernel, which leads to the following objective function (Gosain and Dahiya 2016).

$$J = 2 \sum_{j=1}^n \sum_{i=1}^c \mu_{ij}^m \left(1 - K\left(\vec{x}_j, \vec{v}_i\right)\right), \ \sum_{k=1}^c u_{kj} = 1 \tag{7}$$

where $K\left(\vec{x}_j, \vec{v}_i\right)$ is the kernel, which is defined as follows (Gosain and Dahiya 2016).

$$K\left(\vec{x}_j, \vec{v}_i\right) = 1 - \tanh\left(-\frac{d_{ij}^2}{\omega^2}\right), \ d_{ij}^2 = \left(\vec{x}_j - \vec{v}_i\right)^{\mathrm{T}} A \left(\vec{x}_j - \vec{v}_i\right) \tag{8}$$

where ω is the kernel width and is calculated by the following equations (Gosain and Dahiya 2016).

$$\begin{aligned} \vec{v} &= \frac{1}{n} \sum_{j=1}^n \vec{x}_j \\ \bar{d} &= \frac{1}{n} \sum_{j=1}^n \left\| \vec{x}_j - \vec{v} \right\|_A^2 \\ \omega^2 &= \frac{1}{n-1} \sum_{j=1}^n \left(\left\| \vec{x}_j - \vec{v} \right\|_A^2 - \bar{d} \right)^2 \end{aligned} \tag{9}$$

It is shown that the following cluster centers and Type-2 membership functions minimize (7) (Gosain and Dahiya 2016).

$$\begin{aligned} \mu_{ij} &= \left[\sum_{k=1}^c \left(\frac{1 - K\left(\vec{x}_j, \vec{v}_i\right)}{1 - K\left(\vec{x}_j, \vec{v}_k\right)} \right)^{\frac{1}{m-1}} \right]^{-1}, \ \vec{v}_i \\ &= \frac{\sum_{j=1}^n \mu_{ij}^m K\left(\vec{x}_j, \vec{v}_i\right) \left(2 - K\left(\vec{x}_j, \vec{v}_i\right)\right) \vec{x}_j}{\sum_{j=1}^n \mu_{ij}^m K\left(\vec{x}_j, \vec{v}_i\right) \left(2 - K\left(\vec{x}_j, \vec{v}_i\right)\right)} \end{aligned} \tag{10}$$

3.4 Possibilistic C-means (PCM)

The possibilistic C-means (PCM) algorithm is proposed for data with noise and outliers. Objective function of this algorithm is as follows (Askari 2017b).

$$J = \sum_{j=1}^n \sum_{i=1}^c t_{ij}^\eta d_{ij}^2 + \sum_{i=1}^c \gamma_i \sum_{j=1}^n \left(1 - t_{ij}\right)^\eta \tag{11}$$

where t_{ij} is typicality of $\vec{x}_j$ in the ith cluster and η is the possibilistic exponent. Update equations of the PCM algorithm are as follows (Askari 2017b).

$$\vec{v}_i = \frac{\sum_{j=1}^n t_{ij}^\eta \vec{x}_j}{\sum_{j=1}^n t_{ij}^\eta}, \ t_{ij} = \left(1 + \left(\frac{d_{ij}^2}{\gamma_i}\right)^{\frac{1}{\eta-1}}\right)^{-1}, \ \gamma_i = \frac{\sum_{j=1}^n u_{ij}^m d_{ij}^2}{\sum_{j=1}^n d_{ij}^2} \tag{12}$$

where u_{ij} in γ_i are calculated from (3) by applying the FCM algorithm to the data.

3.5 Possibilistic fuzzy C-means (PFCM) algorithm

The PCM algorithm designed for noisy data suffers from two serious drawbacks including coincident clusters and sensitivity to initialization. The FCM and PCM algorithms are combined to provide the hybrid possibilistic fuzzy C-means (PFCM) algorithm to overcome these shortcomings (Askari 2017b). Objective function of this algorithm is as follows.

$$\begin{aligned} J &= \sum_{j=1}^n \sum_{i=1}^c \left(u_{ij}^m + t_{ij}^\eta\right) d_{ij}^2 + \sum_{i=1}^c \gamma_i \sum_{j=1}^n \left(1 - t_{ij}\right)^\eta, \ \sum_{i=1}^c u_{ij} \\ &= 1 \end{aligned} \tag{13}$$

It is shown that update equations of this algorithm are as follows.

$$\begin{aligned} \vec{v}_i &= \frac{\sum_{j=1}^n \left(u_{ij}^m + t_{ij}^\eta\right) \vec{x}_j}{\sum_{j=1}^n \left(u_{ij}^m + t_{ij}^\eta\right)}, \ u_{ij} = \left[\sum_{k=1}^c \left(\frac{d_{ij}^2}{d_{kj}^2} \right)^{\frac{1}{m-1}} \right]^{-1}, \ t_{ij} \\ &= \left(1 + \left(\frac{d_{ij}^2}{\gamma_i}\right)^{\frac{1}{\eta-1}}\right)^{-1}, \ \gamma_i = \frac{\sum_{j=1}^n u_{ij}^m d_{ij}^2}{\sum_{j=1}^n u_{ij}^m} \end{aligned} \tag{14}$$

3.6 The noise-resistant FCM (nrFCM) algorithm

The present work proposes the noise-resistant FCM (nrFCM) algorithm by replacing the distance d_{ij}^2 in (3) with $f\left(d_{ij}^2\right)$ where f is a differentiable function whose derivative is uniform descending. This function reduces the impacts

of noise and outliers on the cluster centers. Therefore, the following objective function is devised for the nrFCM algorithm.

$$J = \sum_{j=1}^{n} \sum_{i=1}^{c} u_{ij}^{m} f\left(d_{ij}^{2}\right), \quad \sum_{k=1}^{c} u_{kj} = 1 \quad (15)$$

Incorporating the constraint $\sum_{k=1}^{c} u_{kj} = 1$ in the objective function (15) using the Lagrange multipliers results in the following function.

$$J = \sum_{j=1}^{n} \sum_{i=1}^{c} u_{ij}^{m} f\left(d_{ij}^{2}\right) + \sum_{j=1}^{n} \lambda_{j}\left(\sum_{k=1}^{c} u_{kj} - 1\right)$$

Zeroing derivative of this function with respect to cluster centers yields the update equation for $\vec{v}_i$.

$$\frac{\partial J}{\partial \vec{v}_i} = \sum_{j=1}^{n} u_{ij}^{m} f'\left(d_{ij}^{2}\right) \frac{\partial d_{ij}^{2}}{\partial \vec{v}_i} = \sum_{j=1}^{n} u_{ij}^{m} f'\left(d_{ij}^{2}\right)\left(A + A^{T}\right)\left(\vec{x}_j - \vec{v}_i\right)$$
$$= 0 \Rightarrow$$

$$\vec{v}_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} f'\left(d_{ij}^{2}\right) \vec{x}_j}{\sum_{j=1}^{n} u_{ij}^{m} f'\left(d_{ij}^{2}\right)} \quad (16)$$

Similarly, zeroing derivative of this function with respect to u_{ij} results the update equation for the membership grades.

$$\frac{\partial J}{\partial u_{ij}} = m u_{ij}^{m-1} f\left(d_{ij}^{2}\right) + \lambda_{j} = 0 \Rightarrow$$

$$u_{ij} = \left(\frac{-\lambda_j}{m f\left(d_{ij}^{2}\right)}\right)^{\frac{1}{m-1}}, \quad \sum_{k=1}^{c} u_{kj} = 1 \Rightarrow$$

$$u_{ij} = \left[\sum_{k=1}^{c}\left(\frac{f\left(d_{ij}^{2}\right)}{f\left(d_{kj}^{2}\right)}\right)^{\frac{1}{m-1}}\right]^{-1} \quad (17)$$

The nrFCM algorithm converts to the FCM algorithm by choosing $f(x) = x$. Since the cluster centers calculated from (16) depend on f', if the function f is chosen such that f' be uniform descending, the effect of noise and outliers considerably reduces because these points are mostly far away from the cluster centers. Therefore, the following exponential function is used for this purpose.

$$f\left(d_{ij}^{2}\right) = 1 - \exp\left(-\frac{d_{ij}^{2}}{\sigma_i^2}\right), \quad f'\left(d_{ij}^{2}\right) = \frac{1}{\sigma_i^2}\exp\left(-\frac{d_{ij}^{2}}{\sigma_i^2}\right) \quad (18)$$

$$\sigma_i^2 = \frac{\lambda \sum_{j=1}^{n} \mu_{ij}^{m} d_{ij}^{2}}{\sum_{j=1}^{n} \mu_{ij}^{m}}, \quad \mu_{ij} = \left[\sum_{k=1}^{c}\left(\frac{d_{ij}^{2}}{d_{kj}^{2}}\right)^{\frac{1}{m-1}}\right]^{-1} \quad (19)$$

where σ_i is width of the exponential function and λ is a constant (typically $\lambda = 1/9$).

Flowchart of the algorithm is shown in Fig. 2. Degree of fuzziness m, adjusting parameter of clusters width λ, convergence criterion ε, number of clusters c, and the data matrix X are inputs of the algorithm. Since the cluster centers change when the equations are updated, σ_i^2 is
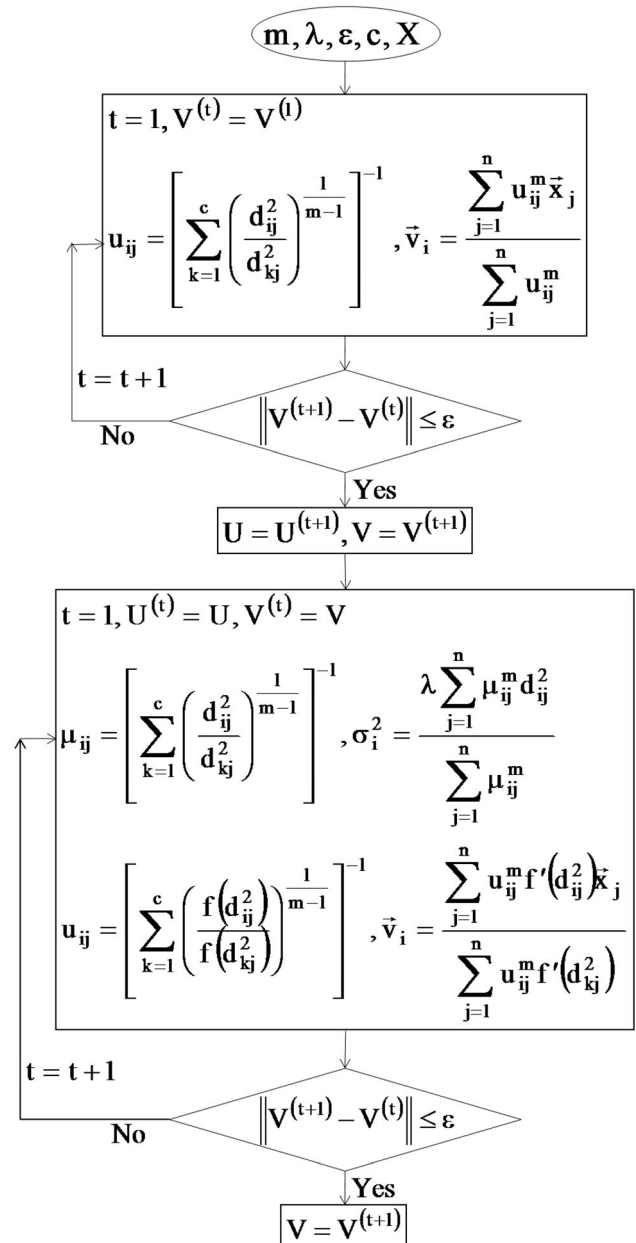

Fig. 2 Flowchart of the nrFCM algorithm

updated in every iteration. When the data are inputted to the algorithm, c data points are randomly chosen from the data matrix X as the initial cluster centers to compute initial cluster centers matrix $V^{(1)}$. Alternatively, a random partition matrix $U = \begin{bmatrix} u_{ij} \end{bmatrix}$ may be generated from which initial cluster centers are calculated using (3). The FCM algorithm starts with these cluster centers and converges when the criterion $\left\| V^{(t+1)} - V^{(t)} \right\| < \varepsilon$ is met that results in the matrix V by which the nrFCM algorithm initializes. Then, the nrFCM algorithm iterates until convergence that results the final cluster centers.

The algorithm is applied through the following step by step procedure.

1. User defined parameters including λ, fuzzy exponent m, termination threshold ε, type of norm matrix A, and number of clusters c are specified. A random partition matrix $U_{c \times n}$ is used to initialize the algorithm where n is number of data points. For example, it can be generated by the command $U = \text{rand}(c, n)$ in MATLAB.
2. The FCM algorithm is then applied to the data. For this purpose, with the initial partition matrix of step 2, initial cluster centers $\vec{v}_i$ are calculated from (3). Membership grades u_{ij} and cluster centers $\vec{v}_i$ are then repeatedly updated using (3) and at each iteration t, the termination criterion $\left\| V^{(t+1)} - V^{(t)} \right\| < \varepsilon$ is examined. When the termination criterion is met, the FCM algorithm stops.
3. The cluster centers calculated by the FCM algorithm are used to initialize the nrFCM algorithm. At each iteration t, μ_{ij} and σ_i are calculated from (19) and then $\vec{v}_i$ and u_{ij} are calculated from (16) and (17), respectively. These values are repeatedly updated until the termination criterion $\left\| V^{(t+1)} - V^{(t)} \right\| < \varepsilon$ is again met.

4 Performance analysis

The main objectives of the fuzzy clustering algorithm are identifying dense regions (clusters) in the data and finding centers of these clusters. Regarding these objectives, the performance of the nrFCM algorithms is compared with those of the FCM, PCM, PFCM, T2FCM, and KT2FCM algorithms in this section. Throughout the article, parameters of the nrFCM algorithm are considered as $\lambda = 1/9$, $m = 2$, $\varepsilon = 0.00001$, $A_{r \times r} = I_r$. For the rest of the algorithms the same values are used for m, ε, and $A_{r \times r}$.

4.1 Calculating actual cluster centers

Actual cluster centers of real data are not known. Therefore, it is not possible to compare the clustering algorithms based on the accuracy of computed cluster centers of real data and instead synthetic data with known cluster centers are used. Three synthetic data sets are used to compare the performance of the FCM, PCM, PFCM, T2FCM, KT2FCM, and nrFCM algorithms. The first data set with three clusters is studied using these algorithms as shown in Fig. 3. The FCM algorithms cannot find actual cluster centers due to noise impacts on the clustering results. As discussed above, the PCM algorithm provides coincident cluster centers especially when the data are highly noisy, which is clearly apparent in Fig. 3.

The FCM and PCM algorithms are combined to derive the hybrid PFCM algorithm for noisy data by utilizing the advantages of both algorithms and avoiding their drawbacks. Although this algorithm performs better than the FCM algorithm for data with few noise and outliers, however, its performance is inferior to that of the FCM algorithm for highly noisy data as shown in Fig. 3. Although cluster centers calculated by the T2FCM algorithm are more accurate than those of the PCM algorithm, but they are too close to each other and far away from the actual cluster centers. However, the KT2FCM algorithm that adds Gaussian kernels to the objective function of the T2FCM algorithm, outperforms the above algorithms. The nrFCM algorithm calculates the most accurate cluster centers by canceling noise impacts by the exponential function as shown in Fig. 3. The actual cluster centers and those calculated by these algorithms are as follows.

$$V_{\text{Actual}} = \begin{bmatrix} -1.0000 & 0.0000 & 1.0000 \\ 0.0000 & 1.0000 & 0.0000 \end{bmatrix},$$

$$V_{\text{FCM}} = \begin{bmatrix} -0.9550 & -0.0004 & 0.9279 \\ 0.3203 & 0.9618 & 0.3104 \end{bmatrix},$$

$$V_{\text{PCM}} = \begin{bmatrix} 0.0509 & 0.0559 & 0.0525 \\ 0.6007 & 0.6085 & 0.6032 \end{bmatrix},$$

$$V_{\text{PFCM}} = \begin{bmatrix} -0.8781 & 0.0155 & 0.8575 \\ 0.3443 & 0.8875 & 0.3303 \end{bmatrix},$$

$$V_{\text{T2FCM}} = \begin{bmatrix} -0.0348 & 0.0401 & 0.1341 \\ 0.6367 & 0.7789 & 0.6409 \end{bmatrix},$$

$$V_{\text{KT2FCM}} = \begin{bmatrix} -0.9598 & 0.0105 & 0.9359 \\ 0.1124 & 0.9429 & 0.1270 \end{bmatrix},$$

$$V_{\text{nrFCM}} = \begin{bmatrix} -1.0019 & -0.0005 & 0.9971 \\ 0.0029 & 1.0005 & 0.0056 \end{bmatrix}$$

To measure the accuracy of different algorithms, the average deviation between the actual and computed cluster
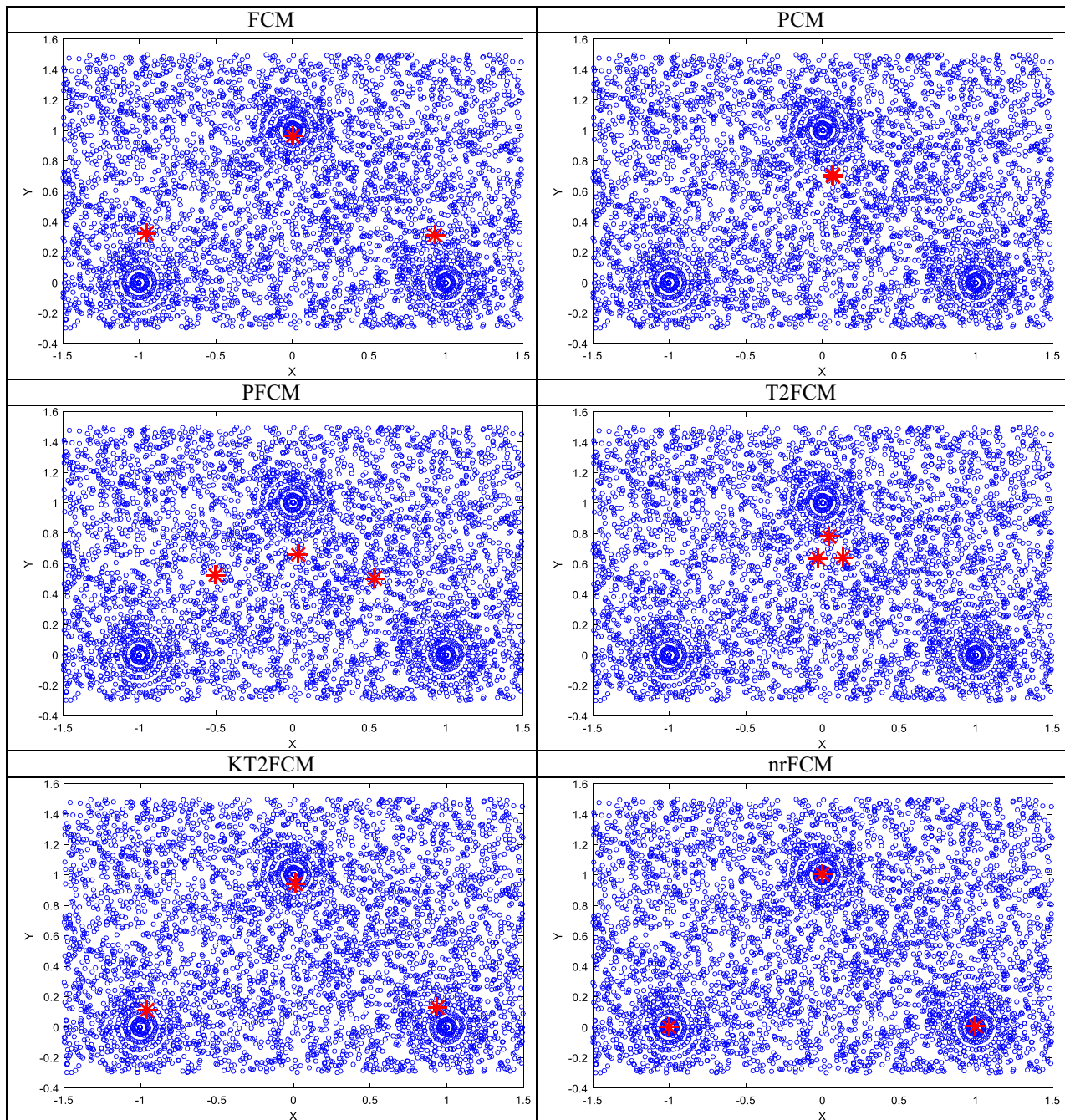
Fig. 3 Clustering the data with three clusters by different algorithms

centers are calculated by the following equation (Askari et al. 2017b). This metric calculates the average distance between the actual and computed cluster centers.

$$E = \frac{1}{c} \sum_{i=1}^{c} \sqrt{\left(\vec{v}_i^* - \vec{v}_i\right)^T A \left(\vec{v}_i^* - \vec{v}_i\right)} \qquad (20)$$

where $\vec{v}_i^*$ and $\vec{v}_i$ are the ith actual and computed cluster centers, respectively, and A is the norm matrix defined in

(2). The following values are calculated for each of the above algorithms.

$E_{\text{FCM}} = 0.2266$, $E_{\text{PCM}} = 0.9097$, $E_{\text{PFCM}} = 0.2795$, $E_{\text{T2FCM}} = 0.8194$, $E_{\text{KT2FCM}} = 0.1066$, $E_{\text{nrFCM}} = 0.0035$

It is observed that the cluster centers calculated by the nrFCM algorithm are the most accurate ones with least deviation from the actual cluster centers.

The second data set with four clusters is studied by these algorithms as shown in Fig. 4. The PFCM algorithm outperforms the FCM algorithm but the PCM and T2FCM algorithms compute coincident clusters. The KT2FCM algorithm is more accurate than the PFCM algorithm. However, the nrFCM algorithm provides the most accurate cluster centers.

The actual cluster centers and those calculated by the algorithms are as follows.

$$V_{\text{Actual}} = \begin{bmatrix} -3.0000 & 0.0000 & 0.0000 & 3.0000 \\ 0.0000 & 7.0000 & 3.0000 & 0.0000 \end{bmatrix},$$

$$V_{\text{FCM}} = \begin{bmatrix} -2.6821 & -0.1260 & 0.0829 & 2.7489 \\ 0.7771 & 7.1040 & 3.5675 & 0.6646 \end{bmatrix},$$

$$V_{\text{PCM}} = \begin{bmatrix} 0.0249 & -0.0468 & 0.0169 & 0.0198 \\ 3.1654 & 6.0873 & 3.1079 & 3.1260 \end{bmatrix},$$

$$V_{\text{PFCM}} = \begin{bmatrix} -2.5035 & -0.0842 & 0.0502 & 2.6660 \\ 0.8423 & 6.8650 & 3.3904 & 0.6494 \end{bmatrix},$$

$$V_{\text{T2FCM}} = \begin{bmatrix} -0.1242 & -0.1242 & 0.2067 & 0.2067 \\ 6.2008 & 6.2008 & 2.0681 & 2.0682 \end{bmatrix}$$

$$V_{\text{KT2FCM}} = \begin{bmatrix} -2.6956 & -0.1185 & 0.0751 & 2.7619 \\ 0.7310 & 7.0838 & 3.5140 & 0.6286 \end{bmatrix},$$

$$V_{\text{nrFCM}} = \begin{bmatrix} -2.9990 & -0.0134 & -0.0023 & 2.9992 \\ -0.0022 & 7.0117 & 3.0055 & 0.0053 \end{bmatrix}$$

Average deviations of the cluster centers computed by different algorithms from the actual cluster centers are as follows that proves the superiority of the proposed algorithm over the others.

$$E_{\text{FCM}} = 0.5717, \ E_{\text{PCM}} = 2.4301, \ E_{\text{PFCM}}$$
$$= 0.5652, \ E_{\text{T2FCM}} = 3.0185, \ E_{\text{KT2FCM}}$$
$$= 0.5322, E_{\text{nrFCM}} = 0.0079$$

When sizes or densities of the clusters are different, cluster centers are misplaced even in noise-free data sets since larger or denser clusters pull cluster centers toward themselves. A noise-free data set of two clusters with different sizes is studied by the above algorithms as shown in Fig. 5. The FCM algorithm identifies center of the larger cluster accurately but center of the smaller one is displaced towards the large cluster. The PCM algorithm does not find the small cluster and computes coincident cluster centers. The PFCM algorithm places both cluster centers within the larger cluster due to the possibilistic term of the algorithm inherited from the PCM algorithm. The T2FCM algorithm performs better than PCM and PFCM algorithms but it is less accurate than FCM algorithm. The KT2FCM algorithm does not realize the presence of the small cluster and places both cluster centers within the large cluster. The nrFCM algorithm that is initialized by the cluster centers calculated by the FCM algorithm, cancels mutual

interactions of the clusters by its exponential function and moves these cluster centers to their actual locations.

4.2 Finding dense regions of data

Quality of clustering is indicated by density. High density shows that the cluster centers calculated by the algorithm are located in dense regions of the data. When there are many data points around each cluster center, a larger density ρ results from the following equation (Askari et al. 2017b).

$$\rho = \frac{\sum_{j=1}^{n} \sum_{i=1}^{c} u_{ij}^{m} d_{ij}^{2}}{\sum_{j=1}^{n} \sum_{i=1}^{c} u_{ij}^{m}} \tag{21}$$

Equation (3) shows that the FCM algorithm updates cluster centers by a linear equation. However, the nrFCM algorithm employs the nonlinear update Eq. (16) because of the exponential function used in the objective function (15). The nonlinear update equation takes more time per iteration and consequently the FCM algorithm is faster than the nrFCM algorithm. These algorithms are applied to 23 synthetic and real data sets and are compared in terms of density of clusters, runtime, and number of iterations as provided in Table 1. Data sets 1 to 4 are the four synthetic noisy data sets studied above and shown in Figs. 1, 3, 4, and 5. Data sets 5 to 17 are real data taken from the UCI Machine Learning Repository. The synthetic noise-free data sets 18 to 23 shown in Fig. 6 contain different numbers of clusters to study performance of the algorithms when there are large numbers of clusters in the data. Since the nrFCM algorithm initializes by the cluster centers calculated by the FCM algorithm and starts when the FCM algorithm converges, this algorithm demands higher number of iterations and runtime. For example, for DATA2 in Table 1, the FCM algorithm converges after 25 iterations with runtime of 7.835 s. Then the nrFCM algorithm starts and terminates after 20 iterations (45–25) and runtime of 15.23 s (23.065–7.835). On average, each iteration of the FCM algorithm takes 0.392 s (7.835/20) and each iteration of the nrFCM algorithm takes 0.761 s (15.23/20). For data sets 1 to 3 (shown in Figs. 1, 3, and 4) that contain noise and outliers, density of clustering results of the nrFCM algorithm is higher than that of the FCM algorithm because the nrFCM algorithm finds actual cluster centers where a large number of data points are concentrated. There is no noise or outliers in data set 4 (shown in Fig. 5) but the FCM algorithm cannot find actual cluster centers because of different sizes and densities of the clusters. Displacement of the cluster centers from their actual locations decreases density of clustering results. There are sporadic observations and probably outliers in the real data sets 5 to 17 and consequently the FCM algorithm cannot find dense
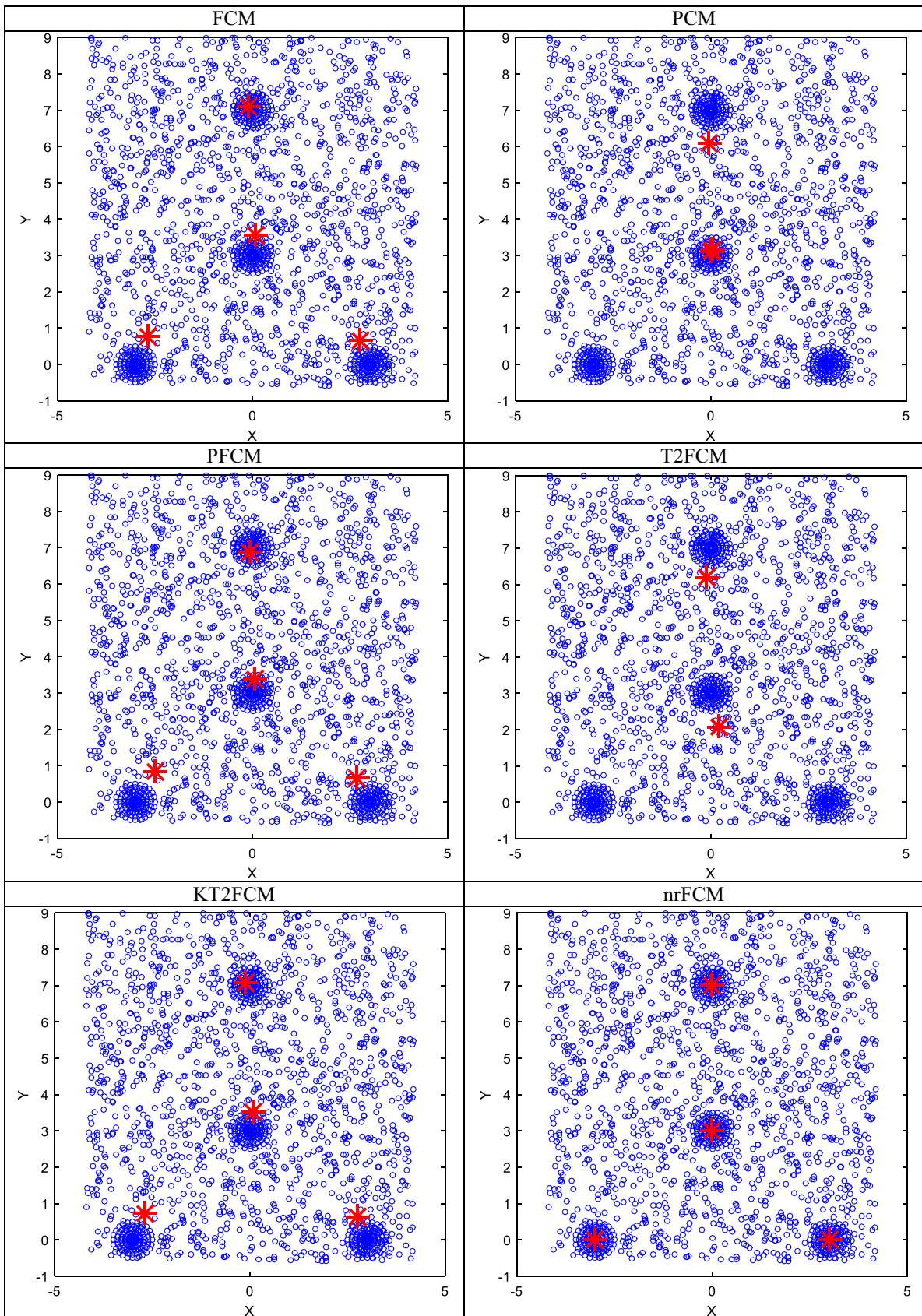
Fig. 4 Clustering the data with four clusters by different algorithms

Fig. 5 Clustering the data with
different cluster sizes and
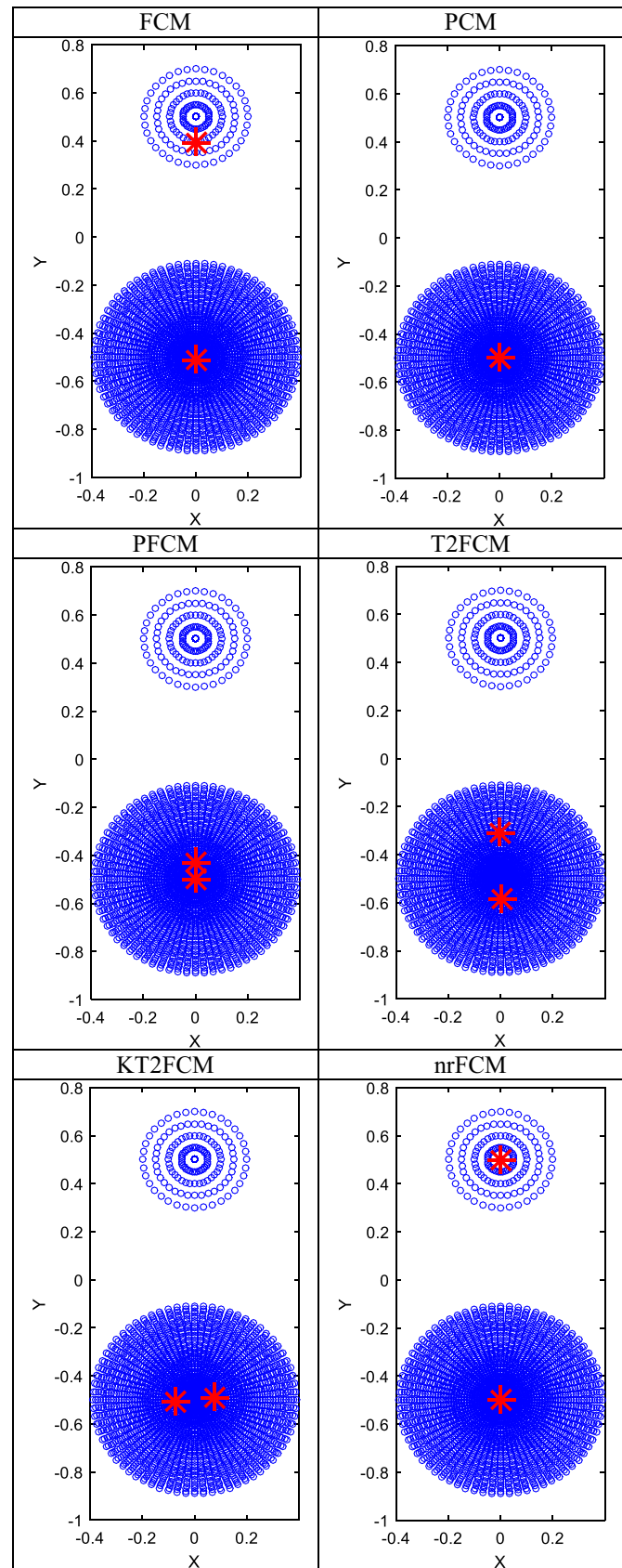densities by various algorithms

Table 1 Number of iterations and runtime (seconds) of the FCM and nrFCM algorithms

|	Data set	c	Density		No. of iterations		Runtime	
			FCM	nrFCM	FCM	nrFCM	FCM	nrFCM
1	DATA1	2	3.259	3.417	11	19	0.018	0.046
2	DATA2	3	0.068	0.078	25	45	7.835	23.065
3	DATA3	4	0.500	0.533	35	48	8.223	15.361
4	DATA4	2	0.022	0.023	55	62	6.176	8.066
5	Climate model simulation crashes	2	0.375	0.548	3	28	0.063	1.305
6	Connectionist bench	2	0.433	0.587	38	58	0.409	0.878
7	Energy efficiency	2	1599	1700	18	33	0.486	1.543
8	Fertility	2	0.572	0.709	19	55	0.103	0.592
9	Glass identification	6	0.122	0.128	25	50	1.312	4.387
10	Haberman survival	2	35.26	40.12	22	98	0.329	2.722
11	Heart disease Cleveland	4	174.35	179.29	77	112	2.978	6.008
12	Ionosphere	2	2.309	2.787	4	40	0.066	1.451
13	IRIS	3	0.134	0.143	26	38	0.358	0.715
14	Pima Indians diabetes	2	2596	3841	46	167	1.359	10.162
15	Seeds	3	0.717	0.748	24	52	0.445	1.559
16	Wine	3	3363	3672	55	149	0.902	4.244
17	Wisconsin prognostic breast cancer	2	10.874	12.479	15	47	0.421	2.608
18	DATA5	5	0.0164	0.0164	16	18	1.902	2.482
19	DATA6	6	0.0136	0.0136	14	16	2.813	3.752
20	DATA7	7	0.0109	0.0109	17	19	5.218	6.654
21	DATA8	8	0.0093	0.0093	17	19	7.833	9.910
22	DATA9	9	0.0079	0.0079	23	25	14.381	17.253
23	DATA10	10	0.0068	0.0068	21	23	17.632	23.450
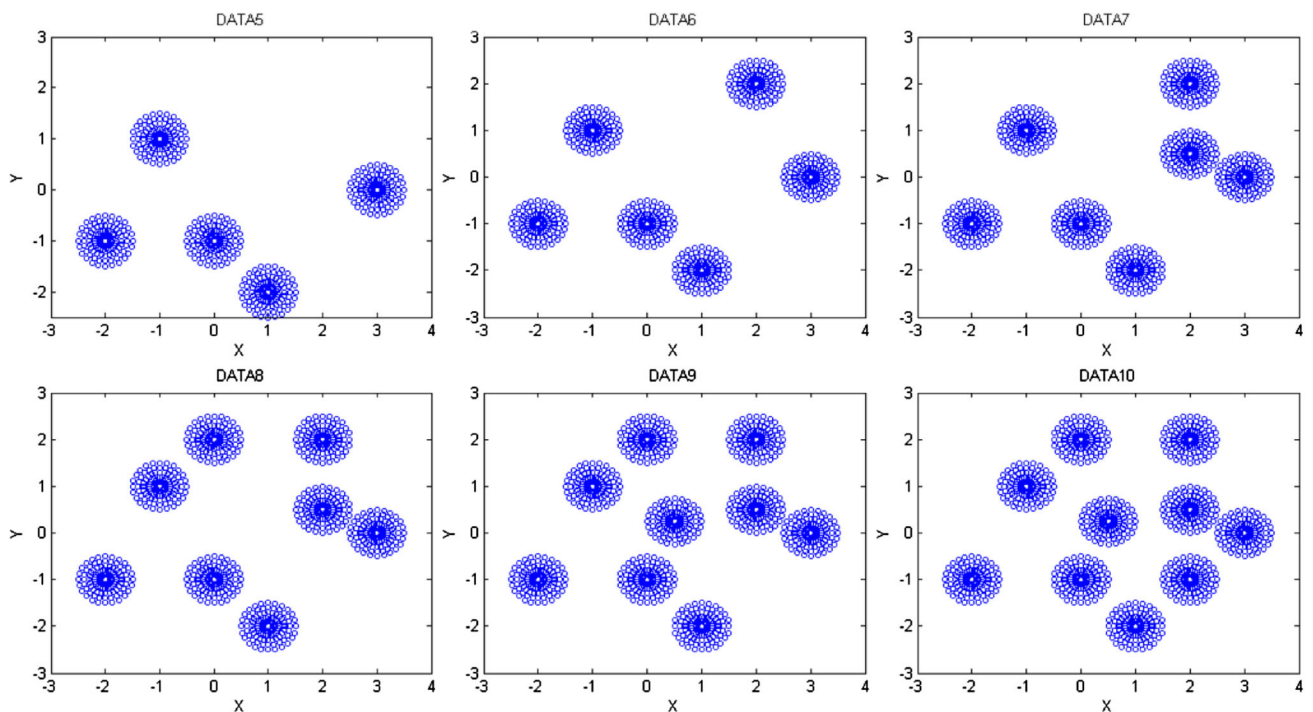
Fig. 6 Six synthetic data sets with no noise and outliers

regions of these data sets but the nrFCM algorithm identifies dense clusters with higher densities as compared to the FCM algorithm. There is no noise and outliers in data sets 18 to 23 and their clusters have the same sizes. Therefore, the FCM algorithm finds actual cluster centers of these data sets and has the same density as the nrFCM algorithm. Since the data sets are clean, the nrFCM algorithm iterates only two times to correct the insignificant displacements of the cluster centers calculated by the FCM algorithm due to the interactions between the clusters. For example, consider the cluster center $\begin{bmatrix} 1 & -2 \end{bmatrix}^T$ of DATA5 shown in Fig. 6. The FCM algorithm calculates this cluster center as $\begin{bmatrix} 1.0030 & -2.0034 \end{bmatrix}^T$ due to the interactions between this cluster and each of the other four clusters and then the nrFCM algorithm moves it to $\begin{bmatrix} 1.0000 & -2.0000 \end{bmatrix}^T$ by damping impacts of other clusters on this cluster by the exponential function.

In summary, according to Table 1, runtime and number of iterations of the nrFCM algorithm is higher than those of the FCM algorithm due to nonlinear update equation for cluster centers. Since the nrFCM algorithm finds dense regions of the data as fuzzy clusters, the density of the clusters calculated by this algorithm is higher than that of the FCM algorithm. Even if there are no noise and outliers in the data, the FCM algorithm cannot find exact cluster centers because of the interactions between the clusters. However, the nrFCM algorithm cancels these interactions by the exponential function introduced in the objective function (15) and finds precise locations of the cluster centers.

5 Conclusions

Presence of noise and outliers in data impairs the accuracy of clustering algorithms and results inaccurate cluster centers. Noise-resistant FCM (nrFCM) algorithm is presented in this article for the data with noise and outliers. The nrFCM algorithm reduces contributions of noise and outliers by introducing an exponential function in the objective function of the FCM algorithm. It is shown that when the data are noisy, the FCM algorithm fails to determine actual cluster centers but the nrFCM algorithm works satisfactorily. Moreover, the nrFCM algorithm finds dense regions of the data more precisely than the FCM algorithm, which is indicated by the density of the clustering results. It is also shown that even if the data are noise-free, the FCM algorithm cannot calculate actual cluster centers precisely because of different cluster sizes and the interactions between clusters.

In summary, the main contribution of the present work is to improve the performance of the well-known FCM algorithm by proposing the nrFCM algorithm with the following characteristics:

1. Finding accurate cluster centers in the presence of noise and outliers.
2. Placing the cluster centers in dense regions of the data so that they represent the actual structure of the data.
3. Finding accurate cluster centers when sizes of the clusters are considerably different.

Despite high accuracy, the nrFCM algorithm incurs higher runtime because of nonlinear update equation for the cluster centers. Therefore, the runtime of the algorithm mainly depends on the function f in Eq. (15) for which exponential function is used in the present work. It is possible to find a proper function capable of canceling noise impacts with less runtime. For example, stepwise linearization of the exponential function may be a proper choice. Moreover, although $\lambda = 1/9$ works properly for the data sets studied in the article, but it is more appropriate to calculate this parameter according to the structure of the data (for example dispersion of the data in different dimensions). This makes the algorithm more robust and efficient.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Ahmed T, Mohamed B, Abdelkader C (2015) Nonlinear system identification using clustering algorithm based on kernel method and particle swarm optimization. Int J Uncertain Fuzziness Knowl-Based Syst 23(5):667–683

Aladag CH, Yolcu U, Egrioglu E, Dalar AZ (2012) A new time invariant fuzzy time series forecasting method based on particle swarm optimization. Appl Soft Comput 12(10):3291–3299

Amezcua J, Melin P (2019) A new fuzzy learning vector quantization method for classification problems based on a granular approach. Granul Comput 4:197–209

Anderson D, Bezdek J, Popescu M, Keller J (2010) Comparing fuzzy, probabilistic, and possibilistic partitions. IEEE Trans Fuzzy Syst 18(5):906–917

Antonelli M, Ducange P, Lazzerini B, Marcelloni F (2016) Multi-objective evolutionary design of granular rule-based classifiers. Granul Comput 1:37–58

Apolloni B, Bassis S, Rota J, Galliani GL, Gioia M, Ferrari L (2016) A neurofuzzy algorithm for learning from complex granules. Granul Comput 1:225–246

Askari S (2017a) A novel and fast MIMO fuzzy inference system based on a class of fuzzy clustering algorithms with interpretability and complexity analysis. Expert Syst Appl 84:301–322

Askari S (2017b) Oil reservoirs classification using fuzzy clustering. Int J Eng 30(9):1391–1400

Askari S, Montazerin N (2015) A high-order multi-variable Fuzzy Time Series forecasting algorithm based on fuzzy clustering. Expert Syst Appl 42(4):2121–2135

Askari S, Montazerin N, Fazel Zarandi MH (2015a) A clustering based forecasting algorithm for multivariable fuzzy time series using linear combinations of independent variables. Appl Soft Comput 35:151–160

Askari S, Montazerin N, Fazel Zarandi MH (2015b) Forecasting semi-dynamic response of natural gas networks to nodal gas consumptions using genetic fuzzy systems. Energy 83:252–266

Askari S, Montazerin N, Fazel Zarandi MH (2016a) High frequency modeling of natural gas networks from low frequency nodal meter readings using time series disaggregation. IEEE Trans Ind Inf 12(1):136–147

Askari S, Montazerin N, Fazel Zarandi MH (2016b) Gas networks simulation from disaggregation of low frequency nodal gas consumption. Energy 112:1286–1298

Askari S, Montazerin N, Fazel Zarandi MH, Hakimi E (2017a) Generalized entropy based possibilistic fuzzy c-means for clustering noisy data and its convergence proof. Neurocomputing 219:186–202

Askari S, Montazerin N, Fazel Zarandi MH (2017b) Generalized possibilistic fuzzy c-means with novel cluster validity indices for clustering noisy data. Appl Soft Comput 53:262–283

Askari S, Montazerin N, Fazel Zarandi MH (2020) Modeling energy flow in natural gas networks using time series disaggregation and fuzzy systems tuned by particle swarm optimization. Appl Soft Comput 92:106332

Aydav PSS, Minz S (2020) Granulation-based self-training for the semi-supervised classification of remote-sensing images. Granular Computing 5:309–327

Beliakov G, Li G, Vu HQ, Wilkin T (2015) Characterizing compactness of geometrical clusters using fuzzy measures. IEEE Trans Fuzzy Syst 23(4):1030–1043

Bezdek JC, Ehrlich R, Full W (1984) FCM: the fuzzy c-means clustering algorithm. Comput Geosci 10(2–3):191–203

Bouzbida M, Hassine L, Chaari A (2017) Robust kernel clustering algorithm for nonlinear system identification. Math Probl Eng 2017:1–11

Chen SM, Chang YC (2010) Multi-variable fuzzy forecasting based on fuzzy clustering and fuzzy rule interpolation techniques. Inf Sci 180(24):4772–4783

Chen SM, Chen SW (2014) Fuzzy forecasting based on two-factors second-order fuzzy-trend logical relationship groups and the probabilities of trends of fuzzy logical relationships. IEEE Trans Cybern 45(3):391–403

Chen SM, Chen SW (2015) Fuzzy forecasting based on two-factors second-order fuzzy-trend logical relationship groups and the probabilities of trends of fuzzy logical relationships. IEEE Trans Cybern 45(3):391–403

Chen SM, Jian WS (2017) Fuzzy forecasting based on two-factors second-order fuzzy-trend logical relationship groups, similarity measures and PSO techniques. Inf Sci 391–392:65–79

Chen SM, Phuong BDH (2017) Fuzzy time series forecasting based on optimal partitions of intervals and optimal weighting vectors. Knowl-Based Syst 118:204–216

Chen SM, Tanuwijaya K (2011a) Fuzzy forecasting based on high-order fuzzy logical relationships and automatic clustering techniques. Expert Syst Appl 38(12):15425–15437

Chen SM, Tanuwijaya K (2011b) Multivariate fuzzy forecasting based on fuzzy time series and automatic clustering techniques. Expert Syst Appl 38(8):10594–10605

Chen SM, Wang NY (2010) Fuzzy forecasting based on fuzzy-trend logical relationship groups. IEEE Trans Syst Man Cybern Part B (Cybern) 40(5):1343–1358

Chen L, Chen CLP, Lu M (2011) A multiple-kernel fuzzy c-means algorithm for image segmentation. IEEE Trans Syst Man Cybern Part B (Cybern) 41(5):1263–1274

Chen SM, Chu HP, Sheu TW (2012) TAIEX forecasting using fuzzy time series and automatically generated weights of multiple factors. IEEE Trans Syst Man Cybern Part A: Syst Hum 42(6):1485–1495

Chen SM, Manalu GMT, Pan JS, Liu HC (2013) Fuzzy forecasting based on two-factors second-order fuzzy-trend logical relationship groups and particle swarm optimization techniques. IEEE Trans Cybern 43(3):1102–1117

Chen SM, Zou XY, Gunawan GC (2019a) Fuzzy time series forecasting based on proportions of intervals and particle swarm optimization techniques. Inf Sci 500:127–139

Chen J, Li Y, Yang X, Zhao S, Zhang Y (2019b) VGHC: a variable granularity hierarchical clustering for community detection. Granul Comput. https://doi.org/10.1007/s41066-019-00195-1

Cheng CH, Cheng GW, Wang JW (2008) Multi-attribute fuzzy time series method based on fuzzy clustering. Expert Syst Appl 34(2):1235–1242

Chintalapudi KK, Kam M (1998) A noise-resistant fuzzy C means algorithm for clustering. IEEE World Congr Comput Intell. https://doi.org/10.1109/FUZZY.1998.686334

Ciucci D (2016) Orthopairs and granular computing. Granular. Computing 1:159–170

Dubois D, Prade H (2016) Bridging gaps between several forms of granular computing. Granul Comput 1:115–126

Duru O, Bulut E (2014) A non-linear clustering method for fuzzy time series: histogram damping partition under the optimized cluster paradox. Appl Soft Comput 24:742–748

Egrioglu E, Aladag CH, Yolcu U, Uslu VR, Erilli NA (2011) Fuzzy time series forecasting method based on Gustafson-Kessel fuzzy clustering. Expert Syst Appl 38(8):10355–10357

Filippone M, Masulli F, Rovetta S (2010) Applying the possibilistic c-means algorithm in kernel-induced spaces. IEEE Trans Fuzzy Syst 18(3):572–584

Gosain A, Dahiya S (2016) Performance analysis of various fuzzy clustering algorithms: a review. Proc Comput Sci 79:100–111

Groll L, Jakel J (2005) A new convergence proof of fuzzy C-means. IEEE Trans Fuzzy Syst 13(3):717–720

Hathaway RJ, Bezdek JC (2001) Fuzzy C-means clustering of incomplete data. IEEE Trans Syst Man Cybern Part B (Cybern) 31(5):735–744

Havens T, Bezdek J, Leckie C, Hall L, Palaniswami M (2012) Fuzzy c-means algorithms for very large data. IEEE Trans Fuzzy Syst 20(6):1130–1146

Kalist V, Ganesan P, Sathish BS, Jenitha JMM, Shaik KB (2015) Possibilistic-fuzzy C-means clustering approach for the segmentation of satellite images in HSL color space. Proc Comput Sci 57:49–56

Kaur P, Gosain A (2010) Density-oriented approach to identify outliers and get noiseless clusters in fuzzy C-means. IEEE Int Conf Fuzzy Syst. https://doi.org/10.1109/FUZZY.2010.5584592

Koutroumbas KD, Xenaki SD, Rontogiannis AA (2018) On the convergence of the sparse possibilistic c-means algorithm. IEEE Trans Fuzzy Syst 26(1):324–337

Krishnapuram R, Keller J (1993) A possibilistic approach to clustering. IEEE Trans Fuzzy Syst 1(2):98–110

Krishnapuram R, Keller J (1996) The possibilistic c-means algorithm: insights and recommendations. IEEE Trans Fuzzy Syst 4(3):385–393

Krisnapuram R, Frigui H, Nasroui O (1995a) Fuzzy and possibilistic shell clustering algorithms and their application to boundary detection and surface approximation-Part I. IEEE Trans Fuzzy Syst 3(1):29–43

Krisnapuram R, Frigui H, Nasroui O (1995b) Fuzzy and possibilistic shell clustering algorithms and their application to boundary detection and surface approximation-Part II. IEEE Trans Fuzzy Syst 3(1):44–60

Lei T, Jia X, Zhang Y, He L, Meng H, Nandi AK (2018) Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering. IEEE Trans Fuzzy Syst 26(5):3027–3041

Li ST, Cheng YC (2010) A stochastic HMM-based forecasting model for fuzzy time series. IEEE Trans Syst Man Cybern Part B (Cybern) 40(5):1255–1266

Lingras P, Haider F, Triff M (2016) Granular meta-clustering based on hierarchical, network, and temporal connections. Granul Comput 1:71–92

Liu H, Cocea M (2017) Granular computing-based approach for classification towards reduction of bias in ensemble learning. Granul Comput 2:131–139

Liu H, Cocea M (2019) Nature-inspired framework of ensemble learning for collaborative classification in granular computing context. Granul Comput 4:715–724

Liu H, Zhang L (2018) Fuzzy rule-based systems for recognition-intensive classification in granular computing context. Granul Comput 3:355–365

Liu Z, Xu S, Zhang Y, Chen CLP (2014) A multiple-feature and multiple-kernel scene segmentation algorithm for humanoid robot. IEEE Trans Cybern 44(11):2232–2240

Livi L, Sadeghian A (2016) Granular computing, computational intelligence, and the analysis of non-geometric input spaces. Granul Comput 1:13–20

Loia V, D'Aniello G, Gaeta A, Orciuoli F (2016) Enforcing situation awareness with granular computing: a systematic overview and new perspectives. Granul Comput 1:127–143

Maji P, Pal SK (2007) Rough set based generalized fuzzy c-means algorithm and quantitative indices. IEEE Trans Syst Man Cybern Part B (Cybern) 37(6):1529–1540

Makrogiannis S, Economou G, Fotopoulos S, Bourbakis NG (2005) Segmentation of color images using multiscale clustering and graph theoretic region synthesis. IEEE Trans Syst Man Cybern Part A: Syst Hum 35(2):224–238

Martino FD, Sessa S (2020) Extended Gustafson-Kessel granular hotspot detection. Granul Comput 5:85–95

Ozdemir D, Akarun L (2011) Fuzzy algorithms for combined quantization and dithering. IEEE Trans Image Process 10(6):923–931

Pal NR, Pal K, Keller JM, Bezdek JC (2005) A possibilistic fuzzy C-means clustering algorithm. IEEE Trans Fuzzy Syst 13(4):517–530

Peters G, Weber R (2016) DCC: a framework for dynamic granular clustering. Granul Comput 1:1–11

Song Y, Zhang G, Lu H, Lu J (2019) A noise-tolerant fuzzy c-means based drift adaptation method for data stream regression. IEEE Int Conf Fuzzy Syst. https://doi.org/10.1109/FUZZ-IEEE.2019.8859005

Srinivasan T, Palanisamy B (2015) Scalable clustering of high-dimensional data technique using SPCM with ant colony optimization intelligence. Sci World J 2015:1–5

Tolias YA, Panas SM (1998) Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions. IEEE Trans Syst Man Cybern Part A: Syst Hum 28(3):359–369

Tung WL, Quek C (2004) Falcon: neural fuzzy control and decision systems using FKP and PFKP clustering algorithms. IEEE Trans Syst Man Cybern Part B (Cybern) 34(1):686–695

Wang G, Yang J, Xu J (2017) Granular computing: from granularity optimization to multi-granularity joint problem solving. Granul Comput 2:105–120

Wang Q, Wang X, Fang C, Yang W (2020) Robust fuzzy c-means clustering algorithm with adaptive spatial & intensity constraint and membership linking for noise image segmentation. Appl Soft Comput 92:106318

Wong WK, Bai E, Chu AWC (2010) Adaptive time-variant models for fuzzy-time-series forecasting. IEEE Trans Syst Man Cybern Part B (Cybern) 40(6):1531–1542

Wu F, Yan S, Smith JS, Zhang B (2019) Deep multiple classifier fusion for traffic scene recognition. Granul Comput. https://doi.org/10.1007/s41066-019-00182-6

Xie Z, Wang S, Zhang DY, Chung FL, Hanbin (2007) Image segmentation using the enhanced possibilistic clustering method. Inf Technol J 6(4):541–546

Zeng S, Chen SM, Teng MO (2019) Fuzzy forecasting based on linear combinations of independent variables, subtractive clustering algorithm and artificial bee colony algorithm. Inf Sci 484:350–366

Zhang Q, Chen Z (2014) A distributed weighted possibilistic C-means algorithm for clustering incomplete big sensor data. Int J Distrib Sens Netw 2014:1–14

Zhang JS, Leung YW (2004) Improved possibilistic C-means clustering algorithms. IEEE Trans Fuzzy Syst 12(2):209–217

Zhang M, Hall LO, Goldgof DB (2002) A generic knowledge-guided image segmentation and labeling system using fuzzy clustering algorithms. IEEE Trans Syst Man Cybern Part B (Cybern) 32(5):571–582

Zhang Y, Huang D, Ji M, Xie F (2011) Image segmentation using PSO and PCM with Mahalanobis distance. Expert Syst Appl 38(7):9036–9040

Zhang Q, Yang LT, Chen Z, Xia F (2017a) A high-order possibilistic c-means algorithm for clustering incomplete multimedia data. IEEE Syst J 11(4):2160–2169

Zhang Q, Yang LT, Chen Z, Li P (2017b) PPHOPCM: privacy preserving high-order possibilistic C-means algorithm for big data clustering with cloud computing. IEEE Trans Big Data. https://doi.org/10.1109/TBDATA.2017.2701816

Zhang C, Gao R, Qin H, Feng X (2019a) Three-way clustering method for incomplete information system based on set-pair analysis. Granul Comput. https://doi.org/10.1007/s41066-019-00197-z

Zhang Q, Yang LT, Castiglione A, Peng ZC (2019b) Secure weighted possibilistic C-means algorithm on cloud for clustering big data. Inf Sci 479:515–525

Zhu L, Chung FL, Wang S (2009) Generalized fuzzy c-means clustering algorithm with improved fuzzy partitions. IEEE Trans Syst Man Cybern Part B (Cybern) 39(3):578–591