PAPER • OPEN ACCESS

Robust Fuzzy Clustering Algorithm Based on Adaptive Neighbors

To cite this article: Kehua Miao *et al* 2021 *J. Phys.: Conf. Ser.* **2025** 012046

View the article online for updates and enhancements.

Robust Fuzzy Clustering Algorithm Based on Adaptive Neighbors

Kehua Miao[1], Jialin Wang[1], Yunlong Gao[1], Chao Cao[2*], Youwei Xie[3] and Peng Gao[3]

[1] Department of Automation, Xiamen University, Faculty of Aeronautics and Astronautics, Xiamen 361102, China

[2] Third Institute of Oceanography, Ministry of Natural Resources, Xiamen 361005, China

[3] College of Information Engineering, Jimei University, Xiamen 361021, China
Email: 23220191151289@stu.xmu.edu.cn

Abstract. In pattern recognition, fuzzy logic is widely used in unsupervised classification or clustering methods. Fuzzy c-means (FCM) clustering algorithm is a typical dynamic clustering algorithm of the fuzzy c-means algorithm based on the error square sum criterion. It introduces fuzzy membership and optimizes the objective function to obtain each sample point for all classes. The membership degree of the center is used to automatically classify the data sample. However, the FCM algorithm is susceptible to the noise points and outliers, and the unbalanced data structure reduces the generalization performance of the algorithm. In this paper, we propose a fuzzy c-means clustering algorithm with adaptive neighbors weight learning. Through adaptive neighborhood robust weight learning, an adaptive weight vector with robustness and sparsity is obtained. During the optimization, we only activate the k samples with the shortest distance to the cluster center and eliminate extreme noise samples to improve the global robustness and sparsity of the algorithm. Finally, empirical analysis verifies the superiority and effectiveness of the proposed clustering method.

Keywords. Fuzzy clustering; fuzzy c-means; adaptive robust weight learning.

1. Introduction

Clustering analysis is a crucial data mining method, and the purpose is to find the cluster structure contained in the data set. In recent years, many clustering algorithms have been proposed. These clustering algorithms can be mainly classified into [1] (1) based on partition clustering class algorithm, (2) clustering algorithm based on density, (3) clustering algorithm based on graph theory, (4) clustering algorithm based on other models. Among them, clustering algorithms based on partition include K-means [2], K-Medoids [3], and CLARANS [4]. The partition-based clustering algorithms mainly select the center of the data clusters in different ways and divide the data clusters according to the distance between each sample point and the center of each data cluster. They have two main clustering strategies: hard and fuzzy clustering scheme. Hard clustering means the membership degree of each data has only two values: 0 and 1, such as K-means. But K-means can misclassify sample points of the same category into different data clusters easily due to the lack of interaction between data clusters. Based on K-means, Dunn [5] proposed the FCM algorithm, the most classic clustering algorithm in fuzzy clustering schemes, which was further extended by Bezdek [6]. Fuzzy clustering distinguishes the difference through the degree of membership. The introduction of fuzzy logic gives FCM the ability to measure

the uncertainty and imprecision of the samples in the data set. It also enables interactions between different data clusters, which is conducive to improve the ability of the algorithm to extract image region boundaries in image segmentation. Therefore, most of the current clustering algorithms for image segmentation are developed based on FCM, such as [7-9].

However, for the fuzzy clustering algorithm, the outliers will have a greater pulling force on the center of each data cluster, making each data cluster center offset, which reduces the ability to extract the essential structure of the data cluster. Moreover, the existence of non-outliers will make the data cluster structure fuzzier, which increases the difficulty of extracting the essential structure of the data cluster. Therefore, Krishnapuram R proposed the possibility c-means clustering algorithm (PCM) [10], which removes the constraint conditions on membership in the FCM and adds slacks on the membership values of the sample points in each data cluster. Under the role of relaxation, PCM does reduce the influence of noise points to a certain extent, but also reduces the direct interaction force of data clusters, which leads to too weak ability to extract the essential structure of data clusters and insufficient robustness of the algorithm. Zarinbal M proposed a fuzzy c-means clustering algorithm based on relative entropy (REFCM) [11] by changing the model solution method to relax the membership constraints. REFCM introduces the relative entropy regularization technology based on the FCM model and relaxes the membership constraint by adjusting the Lagrangian multiplier corresponding to the membership constraint. However, the algorithm does not illustrate how to choose the Lagrangian multipliers of different sample points and judge their degree of membership relaxation in detail.

In response to the problems mentioned above, we propose a robust fuzzy c-means clustering algorithm with adaptive neighborhood weight learning. Through adaptive neighborhood robust weight learning, we obtain adaptive weight vectors with robustness and sparsity. During the optimization period, only the k samples with the shortest distance to the cluster center are activated. The remaining samples, that is, the extreme noise samples, are eliminated to improve the anti-noise ability and effectiveness of the algorithm. The rest of this paper is organized as follows: Section 2 mainly introduces related work and the third section explicitly introduces the proposed robust fuzzy clustering algorithm based on adaptive neighbors. The fourth section discusses the robustness and superiority of this method through related experiments and the conclusion is given in Section 5.

2. Related Work

2.1. Fuzzy C-Means Clustering Algorithm

The fuzzy c-means (FCM) clustering algorithm is a data clustering method based on the fuzzy objective function. The main purpose of FCM is to divide the sample points of the vector space into c subspace according to some distance measure. The result of clustering is characterized by the membership degree of the data to the clustering center, which is expressed by a numerical value. The model is as follows:

$$\begin{cases} \min_{U,V} J_{FCM}(U,V) = \sum_{j=1}^{N}\sum_{i=1}^{C} u_{ij}^{m} \left\| x_j - v_i \right\|^2 \\ s.t.\ 0 \le u_{ij} \le 1, \sum_{i=1}^{C} u_{ij} = 1, 0 < \sum_{j=1}^{N} u_{ij} < n. \end{cases} \tag{1}$$

For problem (1), solving by Lagrangian multiplier method can get:

$$u_{ij} = \sum_{k=1}^{C} \left(\frac{\left\| x_j - v_i \right\|}{\left\| x_j - v_k \right\|} \right)^{\left(-\frac{2}{m-1}\right)}, v_i = \frac{\sum_{j=1}^{N} (u_{ij}^{m} x_j)}{\sum_{j=1}^{N} (u_{ij}^{m})} \tag{2}$$

Among them, x_j represents the *jth* sample point, $x_j \in X$, $X = [x_1, x_2, ..., x_N] \in R^{d \times N}$, v_i represents the cluster center of the *ith* data cluster, $i = 1, 2, ..., C$, u_{ij} denotes the membership degree of x_j belonging to v_i, and d is the dimension of the data set X, N is the number of sample points, C is the number of data clusters, m is the weight index to control the ambiguity, $m > 1$. $\left\| x_j - v_i \right\|^2$ is expressed as the square of the Euclidean distance.

3. The Proposed Method

3.1. Robust Fuzzy Clustering Algorithm Based on Adaptive Neighbors

Based on the adaptive neighborhood robust weight learning framework, we proposed the robust fuzzy c-means clustering algorithm model with adaptive neighbors weight learning. The objective function of our algorithm is as follows:

$$
\begin{cases}
\min_{P,U,V} J(P,U,V) = \sum_{j=1}^{N} p_j \sum_{i=1}^{c} u_{ij}^m \left\| x_j - v_i \right\|^2 + \beta p_j^2 \\
s.t. \sum_{j=1}^{N} p_j = 1, \sum_{i=1}^{c} u_{i,j} = 1, 0 \leq p_j, u_{i,j} \leq 1
\end{cases}
\tag{3}
$$

where p_j represents the adaptive weight coefficients of each sample point, $j = 1, 2, ..., N$, β is the proportional coefficient, and the remaining letters have the same meaning as in equation (1).

3.2. Optimization Algorithm

Firstly, we fix the value of P and V in order to solve U, by introducing the Lagrange multipliers, equation (3) can be rewritten as the following form.

$$
L_1(U, \lambda) = p_j u_{ij}^m \left\| x_j - v_i \right\|_2^2 + \beta p_j^2 + \sum_{j=1}^{N} \lambda_j (\sum_{i=1}^{c} u_{ij} - 1)
\tag{4}
$$

We take the partial derivative of the function with respect to u_{ij}, and set it as zero. Then, we obtain

$$
u_{ij} = \sum_{k=1}^{C} \left(\frac{\left\| x_j - v_i \right\|}{\left\| x_j - v_k \right\|} \right)^{\left(-\frac{2}{m-1} \right)}
\tag{5}
$$

Secondly, we fix the value of P and U to solve V, the equation (3) is equivalent to the following optimization problem.

$$
L_2(V) = \sum_{j=1}^{N} p_j u_{ij}^m \left\| x_j - v_i \right\|_2^2 + \beta p_j^2
\tag{6}
$$

We take the derivative of the function with respect to v_i, and set it as zero. Then, we obtain

$$
v_i = \frac{\sum_{j=1}^{N} (u_{ij}^m x_j p_j)}{\sum_{j=1}^{N} (u_{ij}^m p_j)}
\tag{7}
$$

Thirdly, we fix U and V to solve P and by supposing $g(x_j) = \sum_{i=1}^{C} u_{ij}^m \|x_j - v_i\|_2^2$, the optimization objective function can be transferred into:

$$\begin{cases} \min\limits_{p} J(P,G) = \sum_{j=1}^{N} p_j g(x_j) + \beta p_j^2 \\ s.t. \sum_{j=1}^{N} p_j = 1, (p^T\mathbf{1}=1), 0 \le p_j \le 1, \end{cases} \tag{8}$$

Among them, $g(x_j)$ represents the distance function between the *jth* data point and the center of the data cluster. $p = [p_1, p_2, ..., p_N]^T$ is the weight vector, where p_j is the weight assigned to the *jth* distance term. The first term in equation (8) indicates that the sample points far away from the center of the data cluster should be assigned a small weight, while the second term is a regularization term to avoid trivial solutions and overfitting. An effective technique is used to solve the problem (8) so that the weight vector p has k adaptive neighbors (non-zero entries). That is, only k best neighbor samples are activated. In particular, the following specific derivations are provided to obtain a closed-form solution to the problem (8). Use g_j to represent $g(x_j)$ then problem (8) is equivalent to:

$$\min\limits_{p \ge \mathbf{0}, p^T\mathbf{1}=1} \frac{1}{2}\left\|p + \frac{g}{2\gamma}\right\|_2^2 \tag{9}$$

where $g = [g_1, g_2, ..., g_N]^T$, $\mathbf{0} = [0,...0]^T$, $\mathbf{1} = [1,...,1]^T$, then the Lagrangian function is established as

$$L_3(p,\lambda,\sigma) = \frac{1}{2}\left\|p + \frac{g}{2\gamma}\right\|_2^2 - \lambda(p^T\mathbf{1}-1) - \sigma^T p \tag{10}$$

λ and σ are Lagrangian multipliers. The optimal solution of problem (10) satisfies the KKT conditions, p_j can be summarized as

$$p_j = \left(\lambda - \frac{g_j}{2\gamma}\right)_+ \tag{11}$$

Among them, $(\bullet)_+ = \max(\bullet, 0)$. According to equation (11), p_j is non-negative and is inversely proportional to g_j. In addition, we try to determine λ and γ in equation (11). Therefore, without loss of generality, we assume $g_1 \le g_2 \le ... \le g_N$, therefore, based on the negative correlation between g_j and p_j in formula (11). When only considering the k nearest neighbors of p, we have

$$\begin{cases} p_k > 0 \Rightarrow \lambda - \frac{g_k}{2\gamma} > 0 \\ p_{k+1} = 0 \Rightarrow \lambda - \frac{g_{k+1}}{2\gamma} \le 0 \end{cases} \tag{12}$$

By combining (12) when the constraint $p^T\mathbf{1}=1$, we have

$$\sum_{i=1}^{k}(\lambda - \frac{g_i}{2\gamma}) = 1 \Rightarrow \lambda = \frac{1}{k} + \frac{1}{2\gamma k}\sum_{i=1}^{k} g_i \tag{13}$$

According to the constraints in equation (12) and the result in equation (13), the following inequality can be inferred

$$
\begin{cases} \dfrac{1}{k} > \dfrac{g_k}{2\gamma} - \dfrac{1}{2\gamma k}\sum_{j=1}^{k} g_j \\ \dfrac{1}{k} \le \dfrac{g_{k+1}}{2\gamma} - \dfrac{1}{2\gamma k}\sum_{j=1}^{k} g_j \end{cases} \Rightarrow \dfrac{k}{2} g_k - \dfrac{1}{2}\sum_{j=1}^{k} g_j < \gamma < \dfrac{k}{2} g_{k+1} - \dfrac{1}{2}\sum_{j=1}^{k} g_j \tag{14}
$$

In order to obtain special k non-zero weights, the upper limit $\gamma = \dfrac{k}{2} g_{k+1} - \dfrac{1}{2}\sum_{m=1}^{k} g_m$ is selected. Combining λ and γ in equation (13) and equation (14), p_j in equation (11) can be expressed as:

$$
p_j = (\lambda - \dfrac{g_j}{2\gamma})_+ = (\dfrac{1}{k} + \dfrac{1}{2\gamma k}\sum_{m=1}^{k} g_m - \dfrac{g_j}{2\gamma})_+ = \left(\dfrac{2(\dfrac{k}{2}g_{k+1} - \dfrac{1}{m}\sum_{m=1}^{k}g_m) + \sum_{m=1}^{k}g_m - kg_j}{2k(\dfrac{k}{2}g_{k+1} - \dfrac{1}{2}\sum_{j+1}^{k}g_m)} \right)_+ = \left(\dfrac{g_{k+1} - g_j}{kg_{k+1} - \sum_{m=1}^{k}g_m} \right)_+ \tag{15}
$$

Regarding the weight from equation (15), we find that (1) p_j is non-negative and inversely proportional to g_j, which can improve the local robustness of the algorithm with sample points farther from the center of the data cluster are assigned smaller weights; (2) if $j > k$, then $P_j = 0$, which ensures the sparsity with indicating that only k items with a small distance from the center of the data cluster are considered; (3) k is a manipulable integer parameter that directly controls the number of active samples, which indicates global robustness to outliers. Finally, our algorithm flow is shown in table 1.

Table 1. The proposed algorithm.

| **Pseudo code of ANFCM algorithm** |
|---|
| **Input**: |
| Training set $X \in [x_1, x_2, ..., x_n] \in R^{d \times N}$, the number of nearest neighbors k, the fuzzy index m, the number of data clusters C, the maximum iteration number iteration, and the objective function convergence threshold δ, the initialization iteration number $t = 0$. |
| **Initialization**: cluster center V, the weight vector P, and the objective function J^o. |
| **Iteration start** |
| Step1 Update the membership matrix U according to equation (5). |
| Step2 Updates the data cluster center matrix V according to equation (7). |
| Step3 Updates $G = [g_1, g_{2,...,}g_N]^T$, according to $g(x_j) = \sum_{i=1}^{C} u_{ij}^m \lVert x_j - v_i \rVert_2^2$. |
| Step4 Sort G satisfying $g_1 \le g_2 \le ... \le g_N$. |
| Step5 Updates the weight vector p according to equation (15). |
| Step6 Solve the objective function value J^t according to equation (3). |
| If it satisfies $J^t - J^{t+1} < \delta$ and t < iteration, stop the iteration. Otherwise, $t = t + 1$, and turn to step1. |
| **Output:** |
| Membership matrix U, data cluster center matrix V, weight vector P, and clustering results are obtained. |

4. Experiments

In order to verify the effectiveness and superiority of the algorithm, relevant experimental analysis was performed on the unbalanced data set and UCI data set. At the same time, the practicability of the algorithm was further verified in the image segmentation experiment. In the UCI data set experiment, the rand index is selected to estimate the classification accuracy, and the definition of RI is as follows:

$$RI = \frac{R_{00} + R_{11}}{n(n-1)/2} \tag{16}$$

Among them, R_{00} is the number of pairs of sample points in the same data cluster with the same label, and R_{11} is the number of pairs of sample points in different data clusters with various labels. $RI \in [0,1]$, the larger the RI value, the better the clustering result.

4.1. Unbalanced Data Set

In order to test the generalization performance of the algorithm, a non-spherical data set was created artificially, and FCM was selected as the comparison algorithm to test the generalization ability of the algorithm. Set up two bar-shaped Gaussian distributed data clusters as shown in figure 1, with variances of $\sigma = \begin{bmatrix} 0.01 & 0 \\ 0 & 4 \end{bmatrix}$, changing their center distances of the two data clusters to 3 and 2.2 to determine the robustness of the algorithm in changing the shape of the data cluster. As shown in figure 1, it can be seen that when the center distance of the two data clusters changes sequentially, the ANFCM algorithm can get correct clustering results. In contrast, FCM cannot separate the two categories correctly. The results show that the ANFCM algorithm is not affected by the change of the shape of the data clusters, which means it has better robustness to the data distribution and can still get better clustering results when the data clusters are closely distributed.

| (a) The center distance is 3 | (b) FCM | (c) ANFCM |
| (d) The center distance is 2.2 | (e) FCM | (f) ANFCM |
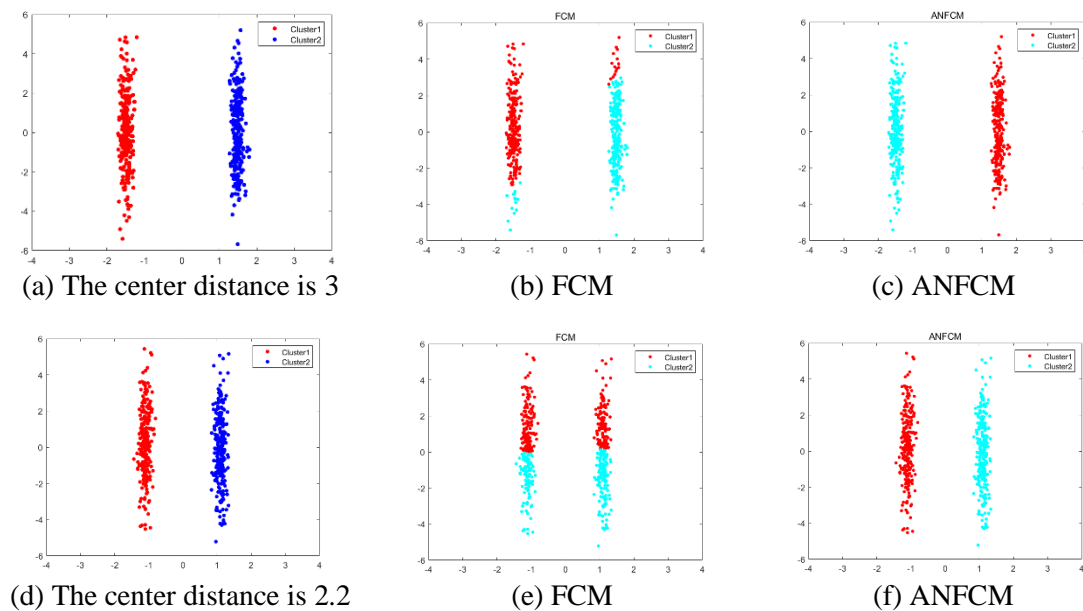
Figure 1. Clustering analysis of ANFCM algorithm on non-spherical data sets.

4.2. UCI Data Set

The UCI database is a commonly used standard test data set for machine learning provided by the University of California, Irvine [12]. Unlike artificial data sets, UCI data sets usually do not meet the various model assumptions of the algorithm on the data set and are typically chosen to measure the overall performance of clustering algorithms. In order to test the effectiveness and accuracy for

processing real data sets, 10 UCI data sets are selected in this paper, FCM, PFCM [13], GIFP-FCM [14-15], csiFCM [16], siibFCM [17], and RBI-FCM [18] as comparison algorithms.

In order to reduce the impact brought by randomness, the experiment adopted the random initialization of the data cluster center. Different algorithms conducted 10 experiments on each data set and recorded the mean and standard deviation of RI value. By observing the experimental results in table 2, it can be found that the ANFCM algorithm can achieve the highest RI value on the 8 data sets in 10 UCI data sets, which fully demonstrates the practicality and effectiveness of the ANFCM algorithm on accurate data.

Table 2. RI index of each algorithm on UCI data sets (%).

| Dataset	FCM	SiibFCM	CSIFCM	GIFPFCM	PFCM	RBIFCM	ANFCM
ecoli	0.7901±0.028	0.7980±0	0.7917±0	0.7487±0.0244	0.7882±0.0015	0.7903±0.0032	**0.8478±0.0077**
breast	0.5196±0	0.5196±0	0.5197±0	0.5147±0.0122	0.5196±0	**0.5433±0.0033**	0.5196±0
zoo	0.8397±0.0274	0.8424±0	0.8439±0	0.2330±0	0.8295±0.0226	0.8385±0.0205	**0.8899±0.0179**
wine	0.7105±0	0.6964±0	0.7136±0	0.6517±0.0569	0.7105±0	0.5128±0.0295	**0.7345±0.0080**
waveform-21	0.6625±0	0.6719±0	0.6575±0	0.6720±0.0248	0.6625±0	0.6625±0	**0.6807±0.0004**
vote	0.6962±0	0.5577±0	0.5248±0	0.6507±0.0928	0.6962±0	0.6962±0	**0.7062±0.0071**
vehicle	0.6506±0	0.6350±0	0.6557±0	0.6298±0.0304	0.6506±0	0.6210±0.0002	**0.6718±0.0003**
segment	**0.8834±0**	0.8739±0	0.8557±0	0.8310±0.0148	0.8802±0.0037	0.8756±0.0018	0.8324±0.0007
pima	0.5499±0	0.5412±0	0.5450±0	0.5457±0.0013	0.5498±0	0.5532±0.0017	**0.5740±0.0067**
balance	0.5910±0.0529	0.5394±0	0.4299±0	0.5360±0.0415	0.5864±0.0376	0.5707±0.0309	**0.6201±0.0052**

4.3. Image Segmentation

Image segmentation is a classic problem in image processing and computer vision, especially in image analysis, understanding and recognition is a crucial technology [19]. Image segmentation divides the image into different regions according to pixel intensity. The pixel intensities in a unified region are similar, and the pixel intensities from different areas have apparent differences. In natural images, many features are difficult to determine. It is crucial to extract useful sites and detect targets through fuzzy clustering methods. Because of the large amount of data in actual image samples, using FCM to cluster extensive sample data will consume a lot of time and space resources. In this paper, three different authentic images are selected. FCM, PCM, and REFCM algorithms are selected as comparative experiments to compare different algorithms' recognition and segmentation effects on image information.

The experimental results are shown in figures 2 and 3. It can be seen that through different clustering algorithms, the original image is divided into different regions. Compared with other algorithms, the ANFCM algorithm can extract the primary information more effectively. The recognition ability of important information in the actual image is more vital, and the detected target obtained is more accurate.

(a)   (b)   (c)   (d)   (e)
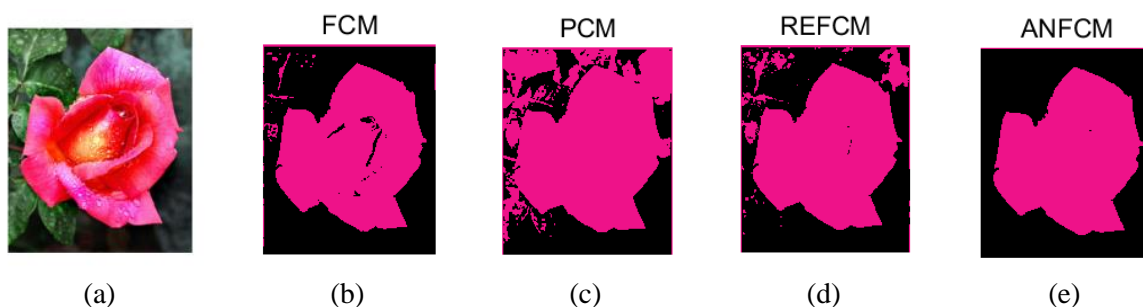
Figure 2. (a) Original picture and the main object of the image obtained by (b) FCM (c) PCM (d) REFCM, and (e) ANFCM methods.
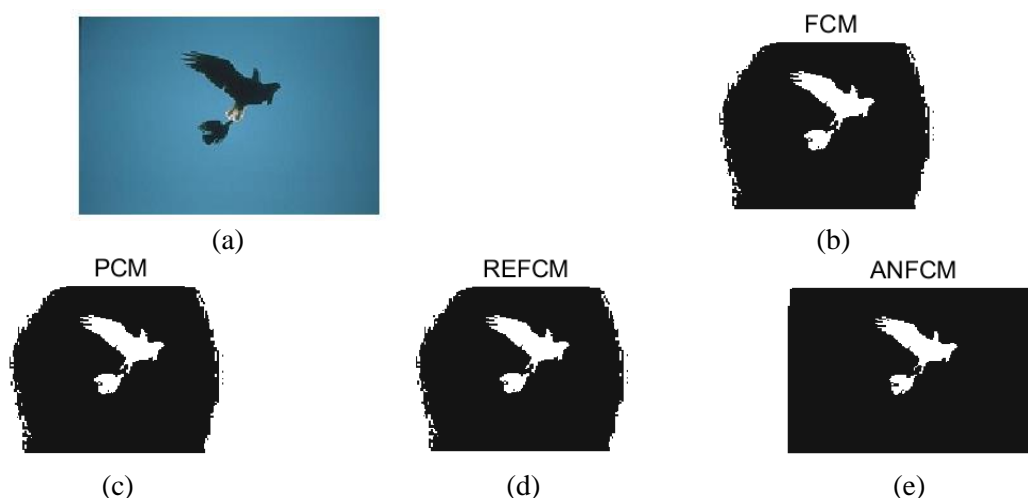
Figure 3. (a) Original picture and the main object of the image obtained by (b) FCM (c) PCM (d) REFCM, and (e) ANFCM methods.

5. Conclusion

This paper proposes a fuzzy c-means clustering algorithm based on a robust weight learning framework of adaptive neighborhoods. Compared with traditional fuzzy clustering algorithms, our algorithm learns robust weights from adaptive neighborhoods in artificial data sets. The proposed models have significant validity and superiority on real data sets. For unbalanced data sets, it has better generalization performance, which can better extract the essential structure of data clusters, and has stronger robust performance and anti-noise ability for noisy data sets. At the same time, the algorithm can also extract the main target more accurately and has better image segmentation performance in image segmentation.

Acknowledgment

References

[1]   Deng Z, Jiang Y, Chung F L, Ishibuchi H, Choi K S and Wang S 2016 Transfer prototype-based fuzzy clustering *IEEE Transactions on Fuzzy Systems* **24** (5) 1210-32.

[2]   Macqueen J B 1967 Some methods for classification and analysis of multivariate observations Proceedings *of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* pp 281-97.

[3]   Park H S and Jun C H 2009 A simple and fast algorithm for K-medoids clustering *Expert Systems with Applications* **36** (2) 3336-41.

[4]   Ng R T and Han J W 2002 CLARANS: A method for clustering objects for spatial data mining *IEEE Transactions on Knowledge and Data Engineering* **14** (5) 1003-16.

[5]   Dunn J C 1973 A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters *Journal of Cybernetics* **3** (3) 32-57.

[6]   Bezdek J C 1981 Pattern recognition with fuzzy objective function algorithms *Advanced Applications in Pattern Recognition* **22** (1171) 203-239.

[7]   Tolias Y A and Panas S M 1998 Image segmentation by a fuzzy clustering algorithm usingadaptive spatially constrained membership functions *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* **28** (3) 359-69.

[8]   Cai W, Chen S and Zhang D 2007 Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation *Pattern Recognition* **40** (3) 825-38.

[9]   Noordam J C, Van Den Broek W H A M and Buydens L M C 2000 Geometrically guided fuzzy c-means clustering for multivariate image segmentation *Proceedings 15th International Conference on Pattern Recognition* pp 462-465.

[10]   Krishnapuram R and Keller J M 1993 A possibilistic approach to clustering *IEEE Transactions on Fuzzy Systems* **1** (2) 98-110.

[11]   Zarinbal M, Fazel Zarandi M H and Turksen I B 2014 Relative entropy fuzzy c-means clustering *Information Sciences* **260** 74-97.

[12]   Merz C J, Blake C L and Keogh E 1998 *UCI Repository of Machine Learning Databases* http://ftp.ics.uci.edu/pub/machine-learning-databases/.

[13]   Pal N R, Pal K, Keller J M and Bezdek J C 2005 A possibilistic fuzzy c-means clustering algorithm *IEEE Transactions on Fuzzy Systems* **13** (4) 517-30.

[14]   Deng Z H, Zhu L and Wang S H 2009 Research on generalized fuzzy c-means clustering algorithm with improved fuzzy partitions *Computer Research and Development* **46** (5) 814-22.

[15]   Zhu L, Chung F L and Wang S 2009 Generalized fuzzy c-means clustering algorithm with improved fuzzy partitions *IEEE Trans. Syst. Man. Cybern. B Cybern.* **39** (3) 578-91.

[16]   Noordam J C 2002 Multivariate image segmentation with cluster size insensitive fuzzy c-means *Chemometrics and Intelligent Laboratory Systems* **64** (1) 65-78.

[17]   Lin P L, Huang P W, Kuo C H and Lai Y H 2014 A size-insensitive integrity-based fuzzy cmeans method for data clustering *Pattern Recognition* **47** (5) 2042-56.

[18]   Wang Z H, Gao Y L and Yang C Y Robust fuzzy c-means clustering algorithmintegrating between-cluster information *Journal of Electronics and Information Technology* **41** (5) 1114-21.

[19]   Borji A, Izadi S and Itti L 2016 ilab-20M: A large-scale controlled object dataset to investigate deep learning *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.