

# Picture fuzzy clustering: a new computational intelligence method

Pham Huy Thong<sup>1</sup> · Le Hoang Son<sup>2,3</sup>

Published online: 20 May 2015  
© Springer-Verlag Berlin Heidelberg 2015

**Abstract** Fuzzy clustering especially fuzzy *C*-means (FCM) is considered as a useful tool in the processes of pattern recognition and knowledge discovery from a database; thus being applied to various crucial, socioeconomic applications. Nevertheless, the clustering quality of FCM is not high since this algorithm is deployed on the basis of the traditional fuzzy sets, which have some limitations in the membership representation, the determination of hesitancy and the vagueness of prototype parameters. Various improvement versions of FCM on some extensions of the traditional fuzzy sets have been proposed to tackle with those limitations. In this paper, we consider another improvement of FCM on the picture fuzzy sets, which is a generalization of the traditional fuzzy sets and the intuitionistic fuzzy sets, and present a novel picture fuzzy clustering algorithm, the so-called FC-PFS. A numerical example on the IRIS dataset is conducted to illustrate the activities of the proposed algorithm. The experimental results on various benchmark datasets of UCI Machine Learning Repository under different scenarios of parameters of the algorithm reveal that FC-PFS has better

clustering quality than some relevant clustering algorithms such as FCM, IFCM, KFCM and KIFCM.

**Keywords** Clustering quality · Fuzzy *C*-means · Intuitionistic fuzzy sets · Picture fuzzy clustering · Picture fuzzy sets

## 1 Introduction

Fuzzy clustering is considered as a useful tool in the processes of pattern recognition and knowledge discovery from a database; thus being applied to various crucial, socioeconomic applications. *Fuzzy C-means* (FCM) algorithm (Bezdek et al. 1984) is a well-known method for fuzzy clustering. It is also considered as a strong aid of rule extraction and data mining from a set of data, in which fuzzy factors are really common and rise up various trends to work on (De Oliveira and Pedrycz 2007; Zimmermann 2001). The growing demands for the exploitation of intelligent and highly autonomous systems put FCM in a great challenge to be applied to various applications such as data analysis, pattern recognition, image segmentation, group-positioning analysis, satellite images and financial analysis like what can be seen nowadays. Nonetheless, the clustering quality of FCM is not high since this algorithm is deployed on the basis of the traditional fuzzy sets, which have some limitations in the membership representation, the determination of hesitancy and the vagueness of prototype parameters. The *motivation of this paper* is to design a novel fuzzy clustering method that could obtain better clustering quality than FCM.

Scanning the literature, we recognize that one of the most popular methods to handle with those limitations is designing improvement versions of FCM on some extensions of the traditional fuzzy sets. Numerous fuzzy clustering algorithms based on the *type-2 fuzzy sets* (T2FS) (Mendel and John

---

Communicated by V. Loia.

---

✉ Le Hoang Son  
lehoangson@tdt.edu.vn  
Pham Huy Thong  
thongph@vnu.edu.vn

- <sup>1</sup> VNU University of Science, Vietnam National University, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam
- <sup>2</sup> Division of Data Science, Ton Duc Thang University, 19 Nguyen Huu Tho, Tan Phong, Ho Chi Minh City, Vietnam
- <sup>3</sup> Faculty of Information Technology, Ton Duc Thang University, 19 Nguyen Huu Tho, Tan Phong, Ho Chi Minh City, Vietnam

2002) were proposed such as in Hwang and Rhee (2007), Linda and Manic (2012), Zarandi et al. (2012) and Ji et al. (2013). Those methods focused on the uncertainty associated with fuzzifier that controls the amount of fuzziness of FCM. Even though their clustering qualities are better than that of FCM, the computational time is quite large so that researchers prefer extending FCM on the *intuitionistic fuzzy sets* (IFS) (Atanassov 1986). Some early researches developing FCM on IFS were conducted by Hung et al. (2004), Iakovidis et al. (2008) and Xu and Wu (2010). Chaira (2011) and Chaira and Panwar (2013) presented another intuitionistic FCM (Chaira's IFCM) method considering a new objective function for clustering the CT scan brain images to detect abnormalities. Some works proposed by Butkiewicz (2012) and Zhao et al. (2013) developed fuzzy features and distance measures to assess the clustering quality. Son et al. (2012a, b, 2013, 2014) and Son (2014a, b, c, 2015) proposed intuitionistic fuzzy clustering algorithms for geodemographic analysis based on recent results regarding IFS and the possibilistic FCM. Kernel-based fuzzy clustering (KFCM) was applied to enhance the clustering quality of FCM such as in Graves and Pedrycz (2010), Kaur et al. (2012) and Lin (2014). Summaries of the recent intuitionistic fuzzy clustering are referenced in (Xu 2012).

Recently, Cuong (2014) have presented *picture fuzzy sets* (PFS), which is a generalization of the traditional fuzzy sets and the intuitionistic fuzzy sets. PFS-based models can be applied to situations that require human opinions involving more answers of types: yes, abstain, no and refusal so that they could give more accurate results for clustering algorithms deployed on PFS. *The contribution* in this paper is a novel picture fuzzy clustering algorithm on PFS, the so-called FC-PFS. Experimental results conducted on the benchmark datasets of UCI Machine Learning Repository are performed to validate the clustering quality of the proposed algorithm in comparison with those of relevant clustering algorithms. The proposed FC-PFS both ameliorates the clustering quality of FCM and enriches the knowledge of developing clustering algorithms on the PFS sets for practical applications. In other words, the findings are significant in both theoretical and practical sides. *The detailed contributions and the rest of the paper* are organized as follows. Section 2 presents the constitution of FC-PFS including,

- Taxonomies of fuzzy clustering algorithms available in the literature that help us understand the developing flow and the reasons why PFS should be used for the clustering introduced in Sect. 2.1. Some basic picture fuzzy operations, picture distance metrics and picture fuzzy relations are also mentioned in this subsection;
- the proposed picture fuzzy model for clustering and its solutions are presented in Sect. 2.2;
- in Sect. 2.3, the proposed algorithm FC-PFS is described.

Section 3 validates the proposed approach through a set of experiments involving benchmark UCI Machine Learning Repository data. Finally, Sect. 4 draws the conclusions and delineates the future research directions.

## 2 Methodology

### 2.1 Taxonomies of fuzzy clustering

Bezdek et al. (1984) proposed the fuzzy clustering problem where the membership degree of a data point  $X_k$  to cluster  $j$ th denoted by the term  $u_{kj}$  was appended to the objective function in Eq. (1). This clearly differentiates with hard clustering and shows that a data point could belong to other clusters depending on the membership degree. Notice that in Eq. (1),  $N$ ,  $C$ ,  $m$  and  $V_j$  are the number of data points, the number of clusters, the fuzzifier and the center of cluster  $j$ th ( $j = 1, \dots, C$ ), respectively.

$$J = \sum_{k=1}^N \sum_{j=1}^C u_{kj}^m \|X_k - V_j\|^2 \rightarrow \min. \quad (1)$$

The constraints for (1) are,

$$\begin{cases} u_{kj} \in [0, 1] \\ \sum_{j=1}^C u_{kj} = 1 \end{cases} \quad (2)$$

Using the Lagrangian method, Bezdek et al. showed an iteration scheme to calculate the membership degrees and centers of the problem (1, 2) as follows.

$$V_j = \frac{\sum_{k=1}^N u_{kj}^m X_k}{\sum_{k=1}^N u_{kj}^m}; \quad j = 1, \dots, C, \quad (3)$$

$$u_{kj} = \frac{1}{\sum_{i=1}^C \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)^{\frac{2}{m-1}}}; \quad k = 1, \dots, N, \quad j = 1, \dots, C. \quad (4)$$

FCM was proven to converge to (local) minimal or the saddle points of the objective function in Eq. (1). Nonetheless, even though FCM is a good clustering algorithm, how to opt suitable number of clusters, fuzzifier, good distance measure and initialization is worth considering since bad selection can yield undesirable clustering results for pattern sets that include noise. For example, in case of pattern sets that contain clusters of different volume or density, it is possible that patterns staying on the left side of a cluster may contribute more for the other than this one. Misleading selection of para-

meters and measures would make FCM fall into local optima and sensitive to noises and outliers.

Graves and Pedrycz (2010) presented a kernel version of the FCM algorithm namely KFCM in which the membership degrees and centers in Eqs. (3, 4) are replaced with those in Eqs. (5, 6) taking into account the kernel distance measure instead of the traditional Euclidean metric. By doing so, the new algorithm is able to discover the clusters having arbitrary shapes such as ring and ‘X’ form. The kernel function used in (5, 6) is the Gaussian expressed in Eq. (7).

$$V_j = \frac{\sum_{k=1}^N u_{kj}^m K(X_k, V_j) X_k}{\sum_{k=1}^N u_{kj}^m K(X_k, V_j)}; \quad j = 1, \dots, C, \tag{5}$$

$$u_{kj} = \frac{1}{\sum_{i=1}^C \left( \frac{1-K(X_k, V_i)}{1-K(X_k, V_j)} \right)^{\frac{1}{m-1}}}; \quad k = 1, \dots, N, \tag{6}$$

$$K(x, y) = \exp(-\|x - y\|^2 / \sigma^2), \quad \sigma > 0. \tag{7}$$

However, modifying FCM itself with new metric measures or new objective functions with penalized terms or new fuzzifier is somehow not sufficient and deploying fuzzy clustering on some extensions of FS such as T2FS and IFS would be a good choice. Hwang and Rhee (2007) suggested deploying FCM on (interval) T2FS sets to handle the limitations of uncertainties and proposed an IT2FCM focusing on fuzzifier controlling the amount of fuzziness in FCM. A T2FS set was defined as follows.

**Definition 1** A type-2 fuzzy set (T2FS) (Mendel and John 2002) in a non-empty set  $X$  is,

$$\tilde{A} = \{(x, u, \mu_{\tilde{A}}(x, u)) | \forall x \in A, \forall u \subseteq J_X \in [0, 1]\}. \tag{8}$$

where  $J_X$  is the subset of  $X$ ,  $\mu_{\tilde{A}}(x, u)$  is the fuzziness of the membership degree  $u(x)$ ,  $\forall x \in X$ . When  $\mu_{\tilde{A}}(x, u) = 1$ ,  $\tilde{A}$  is called the interval T2FS. Similarly, when  $\mu_{\tilde{A}}(x, u) = 0$ ,  $\tilde{A}$  returns to the FS set. The interval type-2 FCM method aimed to minimize both functions below with  $[m_1, m_2]$  is the interval fuzzifier instead of the crisp fuzzifier  $m$  in Eqs. (1, 2).

$$J_1 = \sum_{k=1}^N \sum_{j=1}^C u_{kj}^{m_1} \|X_k - V_j\|^2 \rightarrow \min, \tag{9}$$

$$J_2 = \sum_{k=1}^N \sum_{j=1}^C u_{kj}^{m_2} \|X_k - V_j\|^2 \rightarrow \min. \tag{10}$$

The constraints in (2) are kept intact. By similar techniques to solve the new optimization problem, the interval membership

$U = [\underline{U}, \overline{U}]$  and the crisp centers are calculated in Eqs. (11–13) accordingly. Within these values, after iterations, the objective functions  $J_1$  and  $J_2$  will achieve the minimum.

$$\overline{u}_{kj} = \begin{cases} \frac{1}{\sum_{i=1}^C \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)^{\frac{2}{m_1-1}}} & \text{if } \frac{1}{\sum_{i=1}^C \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)} < \frac{1}{C} \\ \frac{1}{\sum_{i=1}^C \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)^{\frac{2}{m_2-1}}} & \text{otherwise} \end{cases}, \tag{11}$$

$$\underline{u}_{kj} = \begin{cases} \frac{1}{\sum_{i=1}^C \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)^{\frac{2}{m_1-1}}} & \text{if } \frac{1}{\sum_{i=1}^C \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)} \geq \frac{1}{C} \\ \frac{1}{\sum_{i=1}^C \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)^{\frac{2}{m_2-1}}} & \text{otherwise} \end{cases}, \tag{12}$$

$$V_j = \frac{\sum_{k=1}^N \underline{u}_{kj}^m X_k}{\sum_{k=1}^N \underline{u}_{kj}^m}; \quad j = 1, \dots, C, \tag{13}$$

Where  $m$  is a ubiquitous value between  $m_1$  and  $m_2$ . Nonetheless, the limitation of the class of algorithms that deploy FCM on T2FS is heavy computation so that developing FCM on IFS is preferred than FCM on T2FS. IFS (Atanassov 1986), which comprised elements characterized by both membership and non-membership values, is useful mean to describe and deal with vague and uncertain data.

**Definition 2** An intuitionistic fuzzy set (IFS) (Atanassov 1986) in a non-empty set  $X$  is,

$$\hat{A} = \{(x, \mu_{\hat{A}}(x), \gamma_{\hat{A}}(x)) | x \in X\}, \tag{14}$$

where  $\mu_{\hat{A}}(x)$  is the membership degree of each element  $x \in X$  and  $\gamma_{\hat{A}}(x)$  is the non-membership degree satisfying the constraints,

$$\mu_{\hat{A}}(x), \gamma_{\hat{A}}(x) \in [0, 1], \quad \forall x \in X, \tag{15}$$

$$0 \leq \mu_{\hat{A}}(x) + \gamma_{\hat{A}}(x) \leq 1, \quad \forall x \in X. \tag{16}$$

The intuitionistic fuzzy index of an element (also known as the hesitation degree) showing the non-determinacy is denoted as,

$$\pi_{\hat{A}}(x) = 1 - \mu_{\hat{A}}(x) - \gamma_{\hat{A}}(x), \quad \forall x \in X. \tag{17}$$

When  $\pi_{\hat{A}}(x) = 0$ , IFS returns to the FS set. The hesitation degree can be evaluated through the membership function by Yager generating operator (Burillo and Bustince 1996), that is,

$$\pi_{\hat{A}}(x) = 1 - \mu_{\hat{A}}(x) - (1 - \mu_{\hat{A}}(x)^\alpha)^{1/\alpha}, \tag{18}$$

where  $\alpha > 0$  is an exponent coefficient. This operator is used to adapt with the entropy element in the objective function for

intuitionistic fuzzy clustering in Eq. (19) according to Chaira (2011). Most intuitionistic FCM methods, for instance, the IFCM algorithms in Chaira (2011) and Chaira and Panwar (2013) integrated the intuitionistic fuzzy entropy with the objective function of FCM to form the new objective function as in Eq. (19).

$$J = \sum_{k=1}^N \sum_{j=1}^C u_{kj}^m \|X_k - V_j\|^2 + \sum_{j=1}^C \pi_j^* e^{1-\pi_j^*} \rightarrow \min, \quad (19)$$

where,

$$\pi_j^* = \frac{1}{N} \sum_{k=1}^N \pi_{kj}. \quad (20)$$

Notice that when  $\pi_{\hat{A}}(x) = 0$ , the function (19) returns to that of FCM in (1). The constraints for (19–20) are similar to those of FCM so that the authors, for simplicity, separated the objective function in (19) into two parts and used the Lagrangian method to solve the first one and got the solutions as in (3–4). Then, the hesitation degree is calculated through Eq. (18) and used to update the membership degree as follows.

$$u_{kj} = u_{kj} + \pi_{kj}. \quad (21)$$

The new membership degree is used to calculate the centers as in Eq. (3). The algorithm stops when the difference between two consecutive membership degrees is not larger than a pre-defined threshold.

A kernel-based version of IFCM, the so-called KIFCM, has been introduced by Lin (2014). The KIFCM algorithm used Eqs. (5–7) to calculate the membership degrees and the centers under the Gaussian kernel measure. Updating with the hesitation degree is similar to that in IFCM through equation (21). The main activities of KIFCM are analogous to those of IFCM except the kernel function was used instead of the Euclidean distance.

Recently, Cuong (2014) have presented PFS, which is a generalization of FS and IFS. The definition of PFS is stated below.

**Definition 3** A picture fuzzy set (PFS) (Cuong 2014) in a non-empty set  $X$  is,

$$\hat{A} = \{ \langle x, \mu_{\hat{A}}(x), \eta_{\hat{A}}(x), \gamma_{\hat{A}}(x) \rangle \mid x \in X \}, \quad (22)$$

where  $\mu_{\hat{A}}(x)$  is the positive degree of each element  $x \in X$ ,  $\eta_{\hat{A}}(x)$  is the neutral degree and  $\gamma_{\hat{A}}(x)$  is the negative degree satisfying the constraints,

$$\mu_{\hat{A}}(x), \eta_{\hat{A}}(x), \gamma_{\hat{A}}(x) \in [0, 1], \quad \forall x \in X, \quad (23)$$

$$0 \leq \mu_{\hat{A}}(x) + \eta_{\hat{A}}(x) + \gamma_{\hat{A}}(x) \leq 1, \quad \forall x \in X. \quad (24)$$

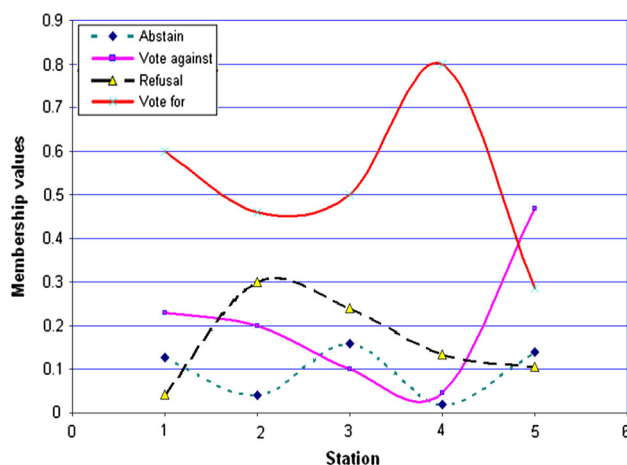


Fig. 1 Picture fuzzy sets

The refusal degree of an element is calculated as  $\xi_{\hat{A}}(x) = 1 - (\mu_{\hat{A}}(x) + \eta_{\hat{A}}(x) + \gamma_{\hat{A}}(x))$ ,  $\forall x \in X$ . In cases  $\xi_{\hat{A}}(x) = 0$  PFS returns to the traditional IFS set. Obviously, it is recognized that PFS is an extension of IFS where the refusal degree is appended to the definition. Yet why we should use PFS and does this set have significant meaning in real-world applications? Let us consider some examples below.

*Example 1* In a democratic election station, the council issues 500 voting papers for a candidate. The voting results are divided into four groups accompanied with the number of papers that are “vote for” (300), “abstain” (64), “vote against” (115) and “refusal of voting” (21). Group “abstain” means that the voting paper is a white paper rejecting both “agree” and “disagree” for the candidate but still takes the vote. Group “refusal of voting” is either invalid voting papers or did not take the vote. This example was happened in reality and IFS could not handle it since the refusal degree (group “refusal of voting”) does not exist.

*Example 2* A patient was given the first emergency aid and diagnosed by four states after examining possible symptoms that are “heart attack”, “uncertain”, “not heart attack”, “appendicitis”. In this case, we also have a PFS set.

From Figs. 1, 2 and 3, we illustrate the PFS, IFS and FS for 5 election stations in Example 1, respectively. We clearly see that PFS is the generalization of IFS and FS so that clustering algorithms deployed on PFS may have better clustering quality than those on IFS and FS. Some properties of PFS operations, the convex combination of PFS, etc., accompanied with proofs are referenced in the article (Cuong 2014).

### 2.2 The proposed model and solutions

In this section, a picture fuzzy model for clustering problem is given. Supposing that there is a dataset  $X$  consisting of  $N$

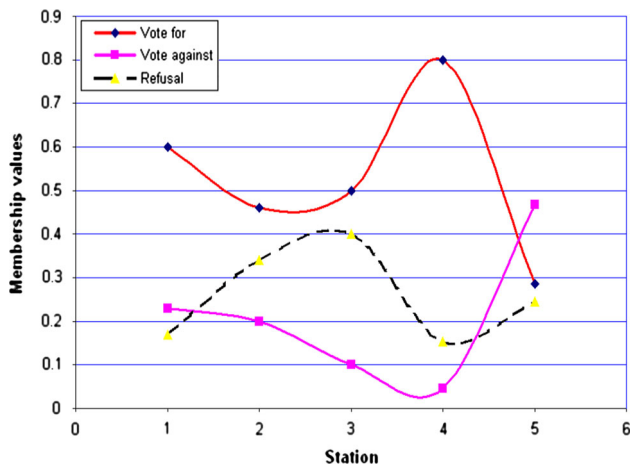


Fig. 2 Intuitionistic fuzzy sets

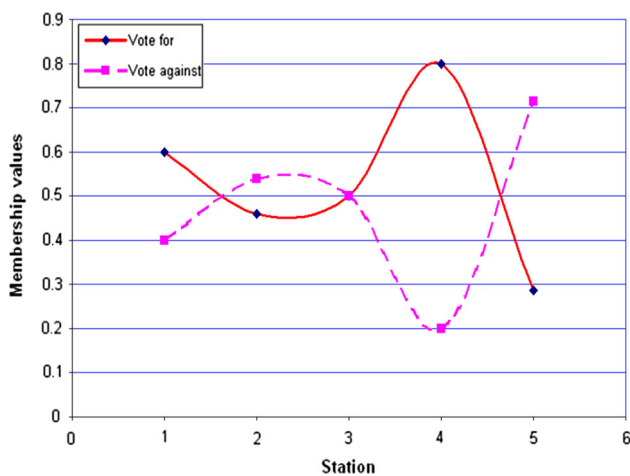


Fig. 3 Fuzzy sets

data points in  $d$  dimensions. Let us divide the dataset into  $C$  groups satisfying the objective function below.

$$J = \sum_{k=1}^N \sum_{j=1}^C (u_{kj}(2 - \xi_{kj}))^m \|X_k - V_j\|^2 + \sum_{k=1}^N \sum_{j=1}^C \eta_{kj}(\log \eta_{kj} + \xi_{kj}) \rightarrow \min, \tag{25}$$

Some constraints are defined as follows:

$$u_{kj} + \eta_{kj} + \xi_{kj} \leq 1, \tag{26}$$

$$\sum_{j=1}^C (u_{kj}(2 - \xi_{kj})) = 1, \tag{27}$$

$$\sum_{j=1}^C \left( \eta_{kj} + \frac{\xi_{kj}}{C} \right) = 1, \tag{28}$$

$$k = 1, \dots, N, \quad j = 1, \dots, C.$$

The proposed model in Eqs. (25–28) relies on the principles of the PFS set. Now, let us summarize the major points of this model as follows.

- The proposed model is the generalization of the intuitionistic fuzzy clustering model in Eqs. (2, 19, 20) since when  $\xi_{kj} = 0$  and the condition (28) does not exist, the proposed model returns to the intuitionistic fuzzy clustering model;
- When  $\eta_{kj} = 0$  and the conditions above are met, the proposed model returns to the fuzzy clustering model in Eqs. (1, 2);
- Equation (27) implies that the “true” membership of a data point  $X_k$  to the center  $V_j$ , denoted by  $u_{kj}(2 - \xi_{kj})$  still satisfies the sum-row constraint of memberships in the traditional fuzzy clustering model.
- Equation (28) guarantees the working on the PFS sets since at least one of two uncertain factors namely the neutral and refusal degrees always exist in the model.
- Another constraint (26) reflects the definition of the PFS sets (Definition 3).

Now, Lagrangian method is used to determine the optimal solutions of model (25–28).

**Theorem 1** *The optimal solutions of the systems (25–28) are:*

$$\xi_{kj} = 1 - (u_{kj} + \eta_{kj}) - (1 - (u_{kj} + \eta_{kj})^\alpha)^{\frac{1}{\alpha}}, \tag{29}$$

$$(k = 1, \dots, N, j = 1, \dots, C),$$

$$u_{kj} = \frac{1}{\sum_{i=1}^C (2 - \xi_{ki}) \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)^{\frac{2}{m-1}}}, \tag{30}$$

$$(k = 1, \dots, N, j = 1, \dots, C),$$

$$\eta_{kj} = \frac{e^{-\xi_{kj}}}{\sum_{i=1}^C e^{-\xi_{ki}}} \left( 1 - \frac{1}{C} \sum_{i=1}^C \xi_{ki} \right), \tag{31}$$

$$j = 1, \dots, C),$$

$$V_j = \frac{\sum_{k=1}^N (u_{kj}(2 - \xi_{kj}))^m X_k}{\sum_{k=1}^N (u_{kj}(2 - \xi_{kj}))^m}, \tag{32}$$

$$(j = 1, \dots, C).$$

*Proof* Taking the derivative of  $J$  by  $v_j$ , we have:

$$\frac{\partial J}{\partial V_j} = \sum_{k=1}^N (u_{kj}(2 - \xi_{kj}))^m (-2X_k + 2V_j), \tag{33}$$

$$(k = 1, \dots, N, j = 1, \dots, C)$$

Since  $\frac{\partial J}{\partial V_j} = 0$ , we have:

$$\sum_{k=1}^N (u_{kj}(2 - \xi_{kj}))^m (-2X_k + 2V_j) = 0, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{34}$$

$$\Leftrightarrow \sum_{k=1}^N (u_{kj}(2 - \xi_{kj}))^m X_k = \sum_{k=1}^N (u_{kj}(2 - \xi_{kj}))^m V_j, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{35}$$

$$\Leftrightarrow V_j = \frac{\sum_{k=1}^N (u_{kj}(2 - \xi_{kj}))^m X_k}{\sum_{k=1}^N (u_{kj}(2 - \xi_{kj}))^m}, \quad (j = 1, \dots, C). \tag{36}$$

The Lagrangian function with respect to  $U$  is,

$$\begin{aligned} L(u) &= \sum_{k=1}^N \sum_{j=1}^C (u_{kj}(2 - \xi_{kj}))^m \|X_k - V_j\|^2 \\ &+ \sum_{k=1}^N \sum_{j=1}^C \eta_{kj} (\log \eta_{kj} + \xi_{kj}) \\ &- \lambda_k \left( \sum_{j=1}^C (u_{kj}(2 - \xi_{kj})) - 1 \right) \end{aligned} \tag{37}$$

Since  $\frac{\partial L(u)}{\partial u_{kj}} = 0$ , we have:

$$\frac{\partial L(u)}{\partial u_{kj}} = m u_{kj}^{m-1} (2 - \xi_{kj})^m \|X_k - V_j\|^2 - \lambda_k (2 - \xi_{kj}) = 0, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{38}$$

$$\Leftrightarrow u_{kj} = \frac{1}{2 - \xi_{kj}} \left( \frac{\lambda_k}{m \|X_k - V_j\|^2} \right)^{\frac{1}{m-1}}, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{39}$$

From Eqs. (37, 49), the solutions of  $U$  are set as follows:

$$\sum_{j=1}^C \left( \frac{\lambda_k}{m \|X_k - V_j\|^2} \right)^{\frac{1}{m-1}} = 1, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{40}$$

$$\Leftrightarrow \lambda_k = \left( \frac{1}{\sum_{j=1}^C \left( m \|X_k - V_j\|^2 \right)^{\frac{1}{1-m}}} \right)^{m-1}, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{41}$$

Plugging (41) into (39), we have:

$$u_{kj} = \frac{1}{\sum_{i=1}^C (2 - \xi_{ki}) \left( \frac{\|X_k - V_j\|}{\|X_k - V_i\|} \right)^{\frac{2}{m-1}}}, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{42}$$

Similarly, the Lagrangian function with respect to  $\eta$  is,

$$\begin{aligned} L(\eta) &= \sum_{k=1}^N \sum_{j=1}^C (u_{kj}(2 - \xi_{kj}))^m \|X_k - V_j\|^2 \\ &+ \sum_{k=1}^N \sum_{j=1}^C \eta_{kj} (\log \eta_{kj} + \xi_{kj}) \\ &- \lambda_k \left( \sum_{j=1}^C \left( \eta_{kj} + \frac{\xi_{kj}}{C} \right) - 1 \right). \end{aligned} \tag{43}$$

$$\frac{\partial L(\eta)}{\partial \eta_{kj}} = \log \eta_{kj} + 1 - \lambda_k + \xi_{kj} = 0, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{44}$$

$$\Leftrightarrow \eta_{kj} = \exp \{ \lambda_k - 1 - \xi_{kj} \}, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{45}$$

From Eqs. (38, 55), we have:

$$\sum_{j=1}^C e^{\lambda_k - 1 - \xi_{kj}} + \frac{1}{C} \sum_{j=1}^C \xi_{kj} = 1, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{46}$$

$$\Leftrightarrow e^{\lambda_k - 1} \sum_{j=1}^C e^{-\xi_{kj}} = 1 - \frac{1}{C} \sum_{j=1}^C \xi_{kj}, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{47}$$

$$\Leftrightarrow e^{\lambda_k - 1} = \frac{1 - \frac{1}{C} \sum_{j=1}^C \xi_{kj}}{\sum_{j=1}^C e^{-\xi_{kj}}}, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{48}$$

Combining (48) with (45), we have:

$$\eta_{kj} = \left( 1 - \frac{1}{C} \sum_{i=1}^C \xi_{ki} \right) \frac{e^{-\xi_{kj}}}{\sum_{i=1}^C e^{-\xi_{ki}}}, \quad (k = 1, \dots, N, j = 1, \dots, C) \tag{49}$$

Finally, using similar techniques of Yager generating operator (Burillo and Bustince 1996), we modify the Eq. (18) by

replacing  $\mu_{\hat{A}}(x)$  by  $(u_{kj} + \eta_{kj})$  to get the value of the refusal degree of an element as follows:

$$\xi_{kj} = 1 - (u_{kj} + \eta_{kj}) - (1 - (u_{kj} + \eta_{kj})^\alpha)^{\frac{1}{\alpha}}, \quad (50)$$

where  $\alpha \in (0, 1]$  is an exponent coefficient used to control the refusal degree in PFS sets. The proof is complete.  $\square$

### 2.3 The FC-PFS algorithm

In this section, the FC-PFS algorithm is presented in details.

---

#### Fuzzy Clustering Method on Picture Fuzzy Sets

---

**I:** Data  $X$  whose number of elements ( $N$ ) in  $d$  dimensions; Number of clusters ( $C$ ); threshold  $\varepsilon$ ; fuzzifier  $m$ ; exponent  $\alpha$  and the maximal number of iteration  $\max Steps > 0$

**O:** Matrices  $u$ ,  $\eta$ ,  $\xi$  and centers  $V$ ;

---

#### FC-PFS Algorithm:

---

- 1:  $t = 0$
  - 2:  $u_{kj}^{(t)} \leftarrow random$ ;  $\eta_{kj}^{(t)} \leftarrow random$ ;  $\xi_{kj}^{(t)} \leftarrow random$  ( $k = \overline{1, N}$ ,  $j = \overline{1, C}$ ) satisfying constraint (36)
  - 3: Repeat
  - 4:  $t = t + 1$
  - 5: Calculate  $V_j^{(t)}$  ( $j = \overline{1, C}$ ) by equation (42)
  - 6: Calculate  $u_{kj}^{(t)}$  ( $k = \overline{1, N}$ ;  $j = \overline{1, C}$ ) by equation (40)
  - 7: Calculate  $\eta_{kj}^{(t)}$  ( $k = \overline{1, N}$ ;  $j = \overline{1, C}$ ) by equation (41)
  - 8: Calculate  $\xi_{kj}^{(t)}$  ( $k = \overline{1, N}$ ;  $j = \overline{1, C}$ ) by equation (39)
  - 9: Until  $\|u^{(t)} - u^{(t-1)}\| + \|\eta^{(t)} - \eta^{(t-1)}\| + \|\xi^{(t)} - \xi^{(t-1)}\| \leq \varepsilon$  or  $t > \max Steps$
- 

## 3 Findings and discussions

### 3.1 Experimental design

In this part, the experimental environments will be described such as,

- *Experimental tools* the proposed algorithm—FC-PFS has been implemented in addition to FCM (Bezdek et al. 1984), IFCM (Chaira 2011), KFCM (Graves and Pedrycz 2010) and KIFCM (Lin 2014) in C programming language and executed them on a Linux Cluster 1350 with eight computing nodes of 51.2GFlops. Each node contains two Intel Xeon dual core 3.2 GHz, 2 GB Ram. The experimental results are taken as the average values after 50 runs.

- *Experimental dataset* the benchmark datasets of UCI Machine Learning Repository such as IRIS, WINE, WDBC (Wisconsin Diagnostic Breast Cancer), GLASS, IONOSPHERE, HABERMAN, HEART and CMC (Contraceptive Method Choice) (University of California 2007). Table 1 gives an overview of those datasets.

**Table 1** The descriptions of experimental datasets

Dataset	No. of elements	No. of attributes	No. of classes	Elements in each classes
IRIS	150	4	3	(50, 50, 50)
WINE	178	13	3	(59, 71, 48)
WDBC	569	30	2	(212, 357)
GLASS	214	9	6	(70, 76, 17, 13, 9, 29)
IONOSPHERE	351	34	2	(126, 225)
HABERMAN	306	3	2	(225, 81)
HEART	270	13	2	(150, 120)
CMC	1473	9	3	(415, 227, 831)

- *Cluster validity measurement* Mean accuracy (MA), the Davies–Bouldin (DB) index (1979) and the Rand index (Vendramin et al. 2010) are used to evaluate the qualities of solutions for clustering algorithms. The DB index is shown as below.

$$DB = \frac{1}{C} \sum_{i=1}^C \left( \max_{j:j \neq i} \left\{ \frac{S_i + S_j}{M_{ij}} \right\} \right), \tag{51}$$

$$S_i = \sqrt{\frac{1}{T_i} \sum_{j=1}^{T_i} |X_j - V_i|^2}, \quad (i = 1, \dots, C), \tag{52}$$

$$M_{ij} = \|V_i - V_j\|, \quad (i, j = 1, \dots, C, i \neq j), \tag{53}$$

where  $T_i$  is the size of cluster  $i$ th.  $S_i$  is a measure of scatter within the cluster, and  $M_{ij}$  is a measure of separation between cluster  $i$ th and  $j$ th. The minimum value indicates the better performance for DB index. The Rand index is defined as,

$$RI = \frac{a + d}{a + b + c + d}, \tag{54}$$

where  $a(b)$  is the number of pairs of data points belonging to the same class in  $R$  and to the same (different) cluster in  $Q$  with  $R$  and  $Q$  being two ubiquitous clusters.  $c(d)$  is the number of pairs of data points belonging to the different class in  $R$  and to the same (different) cluster. The larger the Rand index is, the better the algorithm is.

- *Parameters setting* Some values of parameters such as fuzzifier  $m = 2$ ,  $\varepsilon = 10^{-3}$ ,  $\alpha \in (0, 1)$ ,  $\sigma$  and  $\max \text{Steps} = 1000$  are set up for all algorithms as in Bezdek et al. (1984), Chaira (2011), Graves and Pedrycz (2010) and Lin (2014).
- *Objectives*
  - to illustrate the activities of FC-PFS on a given dataset;
  - to evaluate the clustering qualities of algorithms through validity indices. Some experiments on the computational time of algorithms are also considered;

- to validate the performance of algorithms by various cases of parameters.

### 3.2 An illustration of FC-PFS

First, the activities of the proposed algorithm FC-PFS will be illustrated to classify the IRIS dataset. In this case,  $N = 150$ ,  $r = 4$ ,  $C = 3$ . The initial positive, the neutral and the refusal matrices, whose sizes are  $150 \times 3$ , are initialized as follows:

$$\mu^{(0)} = \begin{pmatrix} 0.174279 & 0.164418 & 0.198673 \\ 0.140933 & 0.169170 & 0.198045 \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ 0.225422 & 0.161006 & 0.125153 \end{pmatrix}, \tag{55}$$

$$\eta^{(0)} = \begin{pmatrix} 0.215510 & 0.321242 & 0.320637 \\ 0.324118 & 0.312415 & 0.330315 \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ 0.306056 & 0.329532 & 0.326154 \end{pmatrix}, \tag{56}$$

$$\xi^{(0)} = \begin{pmatrix} 0.469084 & 0.422466 & 0.402644 \\ 0.433791 & 0.424756 & 0.397048 \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ 0.395095 & 0.419692 & 0.440974 \end{pmatrix}. \tag{57}$$

The distribution of data points according to these initializations is illustrated in Fig. 4. From Step 5 of FC-PFS, the cluster centroids are expressed in Eq. (58).

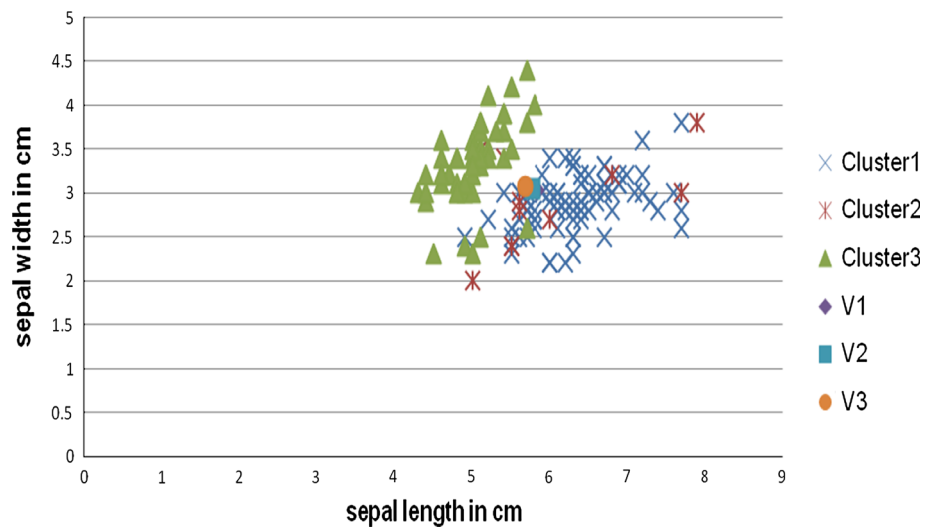
$$V = \begin{pmatrix} 5.833701 & 3.027605 & 3.845183 & 1.248464 \\ 5.784128 & 3.041955 & 3.650875 & 1.148453 \\ 5.677982 & 3.079381 & 3.456761 & 1.073087 \end{pmatrix}. \tag{58}$$

The new positive, neutral and refusal matrices are calculated in the equations below.

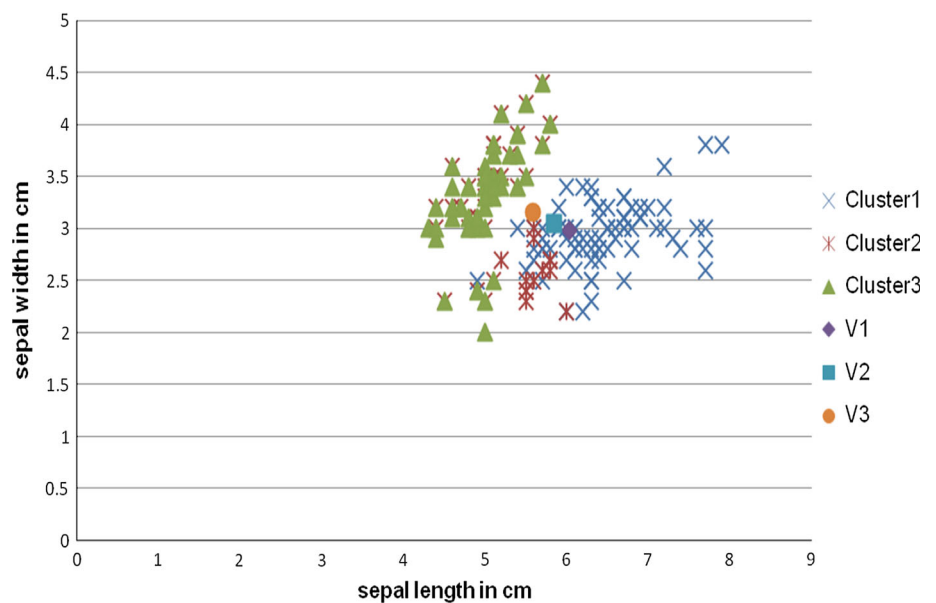
$$\mu^{(1)} = \begin{pmatrix} 0.118427 & 0.169661 & 0.344978 \\ 0.117641 & 0.171776 & 0.340098 \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ 0.400281 & 0.159727 & 0.067458 \end{pmatrix}, \tag{59}$$



**Fig. 4** Clusters in the initialization step



**Fig. 5** Clusters after the first iteration step



$$\eta^{(1)} = \begin{pmatrix} 0.182454 & 0.191161 & 0.194988 \\ 0.190864 & 0.192596 & 0.198008 \\ 0.198376 & 0.193556 & 0.189481 \end{pmatrix}, \tag{60}$$

$$\xi^{(1)} = \begin{pmatrix} 0.495291 & 0.479725 & 0.389716 \\ 0.493854 & 0.478521 & 0.390903 \\ 0.350145 & 0.482186 & 0.499905 \end{pmatrix}. \tag{61}$$

From these matrices, the value of  $\|u^{(t)} - u^{(t-1)}\| + \|\eta^{(t)} - \eta^{(t-1)}\| + \|\xi^{(t)} - \xi^{(t-1)}\|$  is calculated as 0.102, which is larger than  $\epsilon$  so other iteration steps will be made. The distribution of data points after the first iteration step is illustrated in Fig. 5.

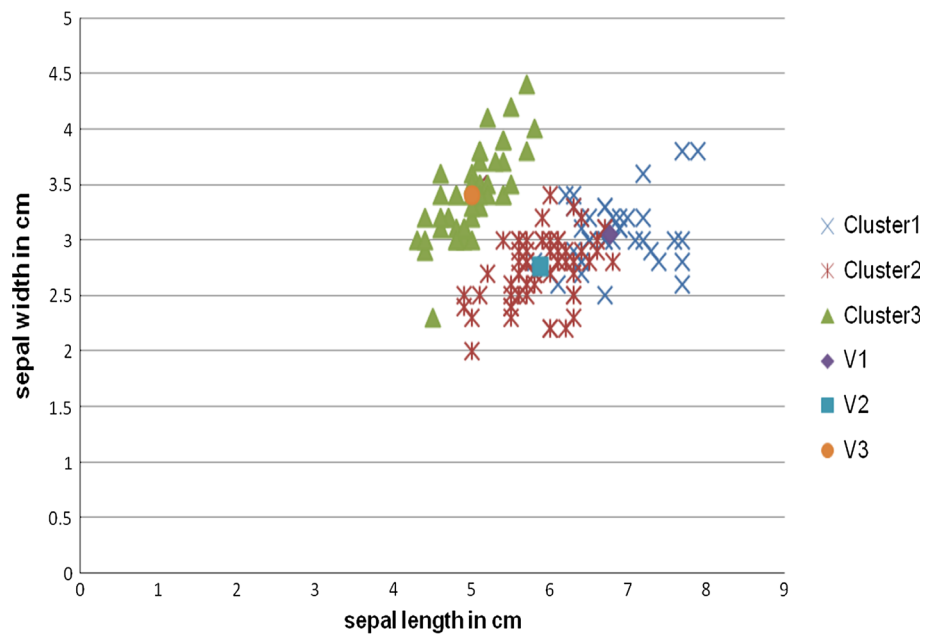
By the similar process, the centers and the positive, neutral and refusal matrices will be continued to be calculated until

the stopping conditions hold. The final positive, neutral and refusal matrices are shown below.

$$\mu^* = \begin{pmatrix} 0.000769 & 0.001656 & 0.551091 \\ 0.004785 & 0.010665 & 0.543119 \\ 0.261915 & 0.350356 & 0.017994 \end{pmatrix}, \tag{62}$$

$$\eta^* = \begin{pmatrix} 0.182155 & 0.182103 & 0.245264 \\ 0.181528 & 0.181223 & 0.242352 \\ 0.186777 & 0.195800 & 0.176763 \end{pmatrix}, \tag{63}$$

$$\xi^* = \begin{pmatrix} 0.489544 & 0.489825 & 0.192064 \\ 0.490654 & 0.492324 & 0.201595 \\ 0.442305 & 0.385736 & 0.493112 \end{pmatrix}. \tag{64}$$

**Fig. 6** Final clusters

The final cluster centroids are expressed in Eq. (65). Final clusters and centers are illustrated in Fig. 6. The total number of iteration steps is 11.

$$V^* = \begin{pmatrix} 6.762615 & 3.048669 & 5.631044 & 2.047229 \\ 5.879107 & 2.757631 & 4.349495 & 1.389834 \\ 5.003538 & 3.403553 & 1.484141 & 0.251154 \end{pmatrix}. \quad (65)$$

### 3.3 The comparison of clustering quality

Second, the clustering qualities and the computational time of all algorithms are validated. The experimental results with the exponent  $\alpha = 0.6$  are shown in Table 2.

It is obvious that FC-PFS obtains better clustering quality than other algorithms in many cases. For example, the Mean Accuracy of FC-PFS for the WINE dataset is 87.1 % which is larger than those of FCM, IFCM, KFCM and KIFCM with the numbers being 85.9, 82.6, 86.2 and 86.6 %, respectively. Similarly, the mean accuracies of FC-PFS, FCM, IFCM, KFCM and KIFCM for the GLASS dataset are 74.5, 71.2, 73.4, 73.5 and 64 %, respectively. When the Rand index of FC-PFS is taken into account, it will be easily recognized that the Rand index of FC-PFS for the CMC dataset is 55.6 % while the values of FCM, IFCM, KFCM and KIFCM are 55.4, 55.1, 50.8 and 48.3 %, respectively. Analogously, the DB value of FC-PFS is better than those of other algorithms. The experimental results on the HEART dataset point out that the DB value of FC-PFS is 2.03, which is smaller and better than those of FCM, IFCM, KFCM and KIFCM with the numbers being 2.05, 2.29, 4.67 and 4.82, respectively. The DB value of FC-PFS on the CMC dataset is equal to that of FCM

and is smaller than those of IFCM, KFCM and KIFCM with the numbers being 2.59, 2.59, 2.85, 4.01 and 3.81, respectively. Taking the average MA value of FC-PFS would give the result 79.85 %, which is the average classification performance of the algorithm. This number is higher than those of FCM (77.3 %), IFCM (77.9 %), KFCM (75.84 %) and KIFCM (70.32 %). Figure 7 clearly depicts this fact.

Nonetheless, the experimental results within a validity index are quite different. For example, the MA values of all algorithms for the IRIS dataset are quite high with the range being (73.3–96.7 %). However, in case of noisy data such as the WDBC dataset, the MA values of all algorithms are small with the range being (56.5–76.2 %). Similar results are conducted for the Rand index where the standard dataset such as IRIS would result in high Rand index range, i.e., (76–95.7 %) and complex, noisy data such as WDBC, IONOSPHERE and HABERMAN would reduce the Rand index ranges, i.e., (51.7–62.5 %), (49.9–52.17 %) and (49.84–49.9 %), respectively. In cases of DB index, the complex data such as GLASS would make high DB values of algorithm than other kinds of datasets. Even though the ranges of validity indices of algorithms are diversified, all algorithms especially FC-PFS would result in high clustering quality with the classification ranges of algorithms being recorded in Table 3.

In Fig. 8, the computational time of all algorithms is depicted. It is clear that the proposed algorithm—FC-PFS is little slower than FCM and IFCM and is faster than KFCM and KIFCM. For example, the computational time of FC-PFS to classify the IRIS dataset is 0.033 seconds (s) in 11 iteration steps. The computational time of FCM, IFCM, KFCM and KIFCM on this dataset is 0.011, 0.01, 0.282 and 0.481 s, respectively. The maximal difference in term of computational time between FC-PFS and other algorithms is occurred

**Table 2** The comparison of algorithms ( $\alpha = 0.6$ )

Data	FC-PFS	FCM	IFCM	KFCM	KIFCM
Mean accuracy (MA) (%)					
IRIS	96.1	96.7 <sup>a</sup>	95.3	84.50	73.3
WINE	87.1 <sup>a</sup>	85.9	82.6	86.20	86.6
WDBC	67.2	62.5	56.5	76.20 <sup>a</sup>	70.6
GLASS	74.5 <sup>a</sup>	71.2	73.4	73.50	64
IONOSPHERE	72.9	71.5	78.6 <sup>a</sup>	60.12	57.6
HABERMAN	76.9	72.2	78.4 <sup>a</sup>	76.70	76.9
HEART	87	85.9	87.7	88.10 <sup>a</sup>	74.9
CMC	77.1 <sup>a</sup>	72.8	70.4	61.40	58.7
Rand index (%)					
IRIS	92.4	95.7 <sup>a</sup>	88.6	84.2	76
WINE	71.3	72.8	73.4	84.2 <sup>a</sup>	68.4
WDBC	56.9	60.2	62.5 <sup>a</sup>	54.1	51.7
GLASS	71.4	70.5	71.9 <sup>a</sup>	67.9	64.7
IONOSPHERE	49.9	50.1	50.2	52.17 <sup>a</sup>	52.1
HABERMAN	49.84	49.87	49.86	49.9 <sup>a</sup>	49.8
HEART	53.9	60.7 <sup>a</sup>	60.7 <sup>a</sup>	50.7	50.8
CMC	55.6 <sup>a</sup>	55.4	55.1	50.8	48.3
DB					
IRIS	3.05	2.91 <sup>a</sup>	3.15	4.21	4.9
WINE	2.77 <sup>a</sup>	2.93	2.88	3.75	3.76
WDBC	1.76 <sup>a</sup>	1.87	1.97	5.03	10.6
GLASS	8.13	5.13 <sup>a</sup>	9.68	13.94	5.9
IONOSPHERE	2.9	2.53 <sup>a</sup>	4.15	2.68	3.9
HABERMAN	2.28	2.21 <sup>a</sup>	2.58	4.92	4.81
HEART	2.03 <sup>a</sup>	2.05	2.29	4.67	4.82
CMC	2.59 <sup>a</sup>	2.59 <sup>a</sup>	2.85	4.01	3.81
Computational time (s)					
IRIS	0.033	0.011	0.01 <sup>a</sup>	0.282	0.481
WINE	0.112 <sup>a</sup>	0.242	0.15	2.56	1.28
WDBC	0.241 <sup>a</sup>	0.323	0.254	3.54	5.007
GLASS	0.534	0.182 <sup>a</sup>	0.277	3.55	4.7
IONOSPHERE	0.177	0.103 <sup>a</sup>	0.194	4.42	4.84
HABERMAN	0.029	0.02 <sup>a</sup>	0.02 <sup>a</sup>	0.126	0.381
HEART	0.071	0.048 <sup>a</sup>	0.07	0.905	0.903
CMC	0.534	0.433	0.422 <sup>a</sup>	6.28	8.15

<sup>a</sup> Best results

on the CMC dataset with the standard deviation being approximately 7.6 s. In some cases of datasets, FC-PFS runs faster than other algorithms, e.g., in the WINE dataset the computational time of FC-PFS, FCM, IFCM, KFCM and KIFCM is 0.112, 0.242, 0.15, 2.56 and 1.28 s, respectively.

Some remarks are the following:

- FC-PFS obtains better clustering quality than other algorithms in many cases;
- The average MA, Rand index and DB values of FC-PFS in this experiment are 79.85, 62.7 and 3.19 %, respectively;

- FC-PFS takes small computational time (approx. 0.22 s) to classify a dataset.

### 3.4 The validation of algorithms by various parameters

Third, we validate whether or not the change of exponent  $\alpha$  has great impact to the performance of algorithms. To understand the influence of this matter, statistics of the numbers of times that all algorithms obtain best results have been made in Table 4.

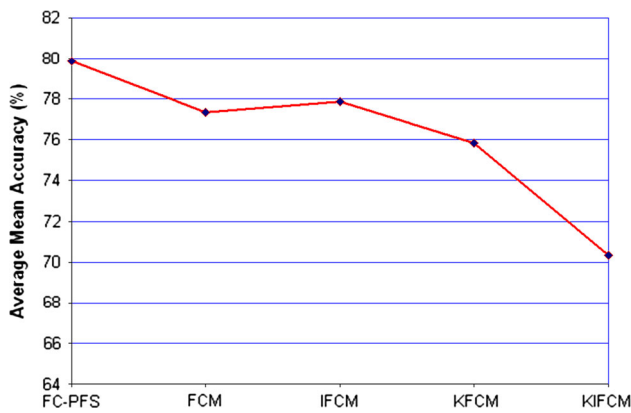


Fig. 7 The average accuracies of algorithms

Table 3 The classification ranges of algorithms

Algorithms	MA (%)	Rand index (%)	DB
FC-PFS	67.2–96.1	49.84 –92.4	1.76–8.13
FCM	62.5–96.7	49.87–95.7	1.87–5.13
IFCM	56.5–95.3	49.86–88.6	1.97–9.68
KFCM	60.12–88.1	49.9–84.2	2.68–13.94
KIFCM	57.6–86.6	48.3–76	3.76–10.6

This table states that when the value of the exponent is small, the number of times that FC-PFS obtains best Mean Accuracy (MA) values among all algorithms is small, e.g., 2 times with  $\alpha = 0.2$ . In this case, FC-PFS is not as effective as FCM since this algorithm achieves 3 times “Best results”. However, when the value of the exponent increases, the clustering quality of FC-PFS is also getting better. The numbers

of times that FC-PFS obtains best MA values among all when  $\alpha = 0.4$ ,  $\alpha = 0.6$  and  $\alpha = 0.8$  are 2, 3 and 3, respectively. Thus, large value of exponent should be chosen to achieve high clustering quality of FC-PFS.

In Fig. 9, the average mean accuracies of algorithms will be computed for a given exponent and list those average MA values in the chart. It has been revealed that FC-PFS is quite stable with the clustering quality expressed through the MA value being approximately 80.6 %, while those of FCM, IFCM, KFCM and KIFCM being 77.3, 72.5, 75.8 and 67.4 %, respectively. The clustering quality of FC-PFS was proven to be better than those of other algorithms through various cases of exponents. In Fig. 10, the average computational time of algorithms will be depicted by exponents. It has been shown that the computational time of FC-PFS is larger than those of FCM and IFCM but smaller than those of KFCM and KIFCM. The average computational time of FC-PFS to cluster a given dataset is around 0.22 s.

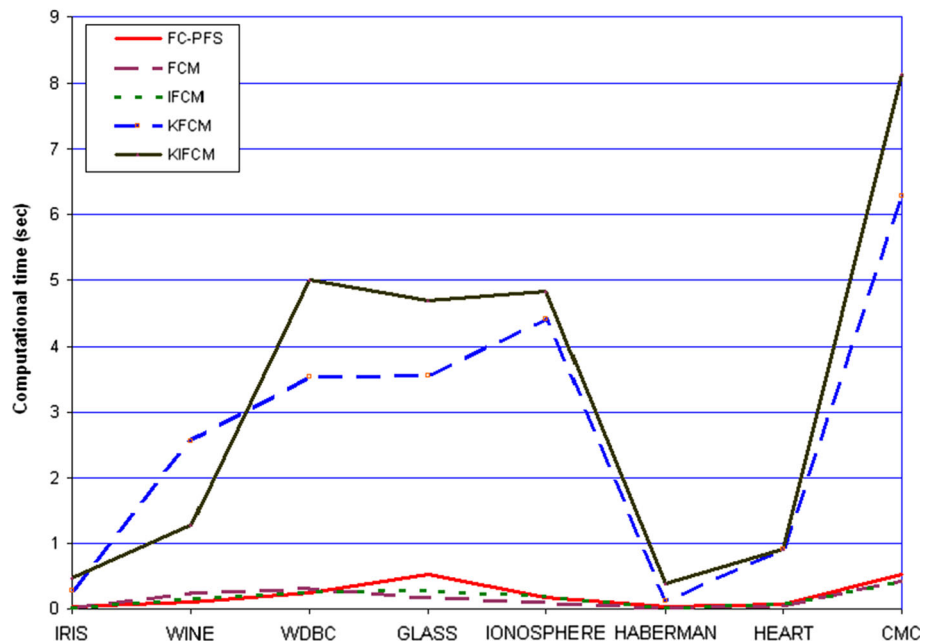
Some remarks are the following:

- The value of exponent should be large to obtain the best clustering quality in FC-PFS;
- the clustering quality of FC-PFS is stable through various cases of exponents (approx. MA ~ 80.6 %). Furthermore, it is better than those of FCM, IFCM, KFCM and KIFCM.

### 4 Conclusions

In this paper, we aimed to enhance the clustering quality of fuzzy C-means (FCM) and proposed a novel picture fuzzy clustering algorithm on picture fuzzy sets, which are gener-

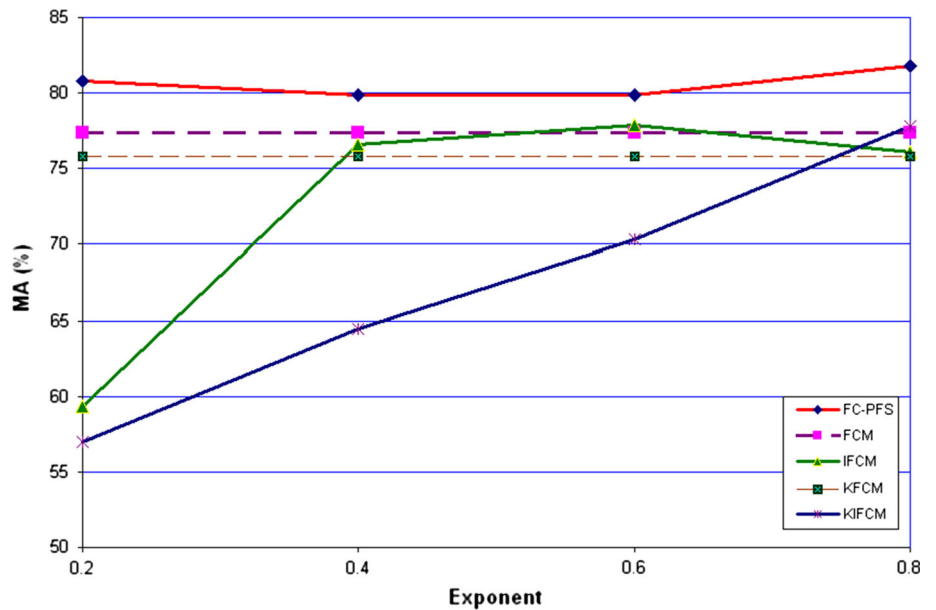
Fig. 8 The computational time of algorithms



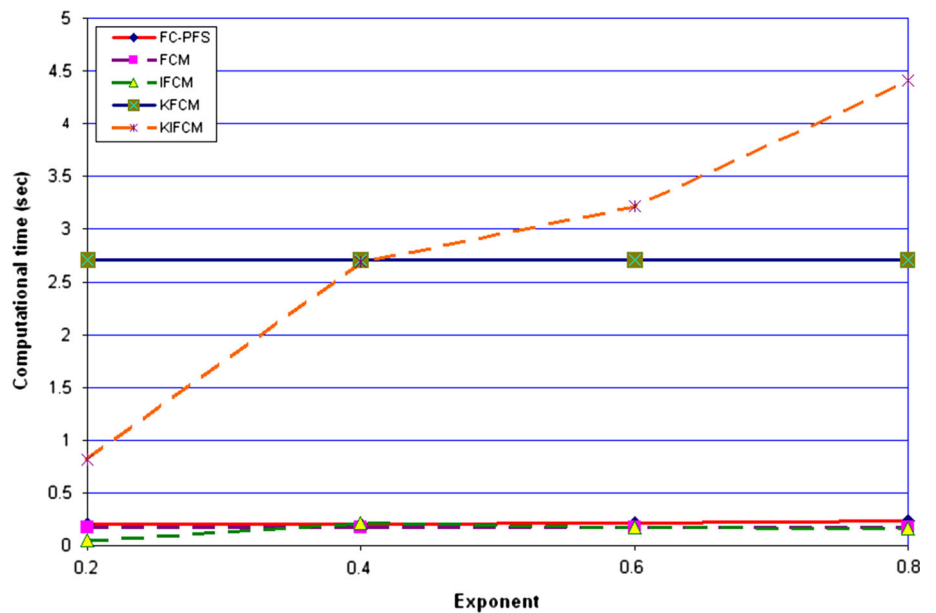
**Table 4** The statistics of best results

Algorithms	MA				Rand index				DB			
	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$
FC-PFS	2	2	3	3	2	2	1	1	4	4	4	2
FCM	1	2	1	0	4	4	2	3	3	4	5	4
IFCM	1	2	2	1	2	1	3	5	0	0	0	2
KFCM	3	1	2	1	1	2	3	2	0	0	0	0
KIFCM	1	1	0	3	0	1	0	1	1	0	0	1

**Fig. 9** The mean accuracies of algorithms by exponents



**Fig. 10** The computational time of algorithms by exponents (s)



alized from the traditional fuzzy sets and intuitionistic fuzzy sets. A detailed description of the previous works regarding the fuzzy clustering on the fuzzy sets and intuitionistic fuzzy

sets was presented to facilitate the motivation and the mechanism of the new algorithm. By incorporating components of the PFS set into the clustering model, the proposed algorithm

produced better clustering quality than other relevant algorithms such as fuzzy *C*-means (FCM), intuitionistic FCM (IFCM), kernel FCM (KFCM) and kernel IFCM (KIFCM). Experimental results conducted on the benchmark datasets of UCI Machine Learning Repository have re-confirmed this fact even in the case of the values of exponents changed and showed the effectiveness of the proposed algorithm. Further works of this theme aim to modify this algorithm in distributed environments and apply it to some forecast applications such as stock prediction and weather nowcasting.

**Acknowledgments** This work is sponsored by a Vietnam National University Scientist Links project, entitled: “To promote fundamental research in the field of natural sciences and life, social sciences and humanities, science of engineering and technology, interdisciplinary science” under the Grant Number QKHCN.15.01.

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Atanassov KT (1986) Intuitionistic fuzzy sets. *Fuzzy Sets Syst* 20:87–96
- Bezdek JC, Ehrlich R, Full W (1984) FCM: the fuzzy *c*-means clustering algorithm. *Comput Geosci* 10(2):191–203
- Burillo P, Bustince H (1996) Entropy on intuitionistic fuzzy set and on interval-valued fuzzy set. *Fuzzy Sets Syst* 78:305–316
- Butkiewicz BS (2012) Fuzzy clustering of intuitionistic fuzzy data. In: Rutkowski L, Korytkowski M, Scherer R, Tadeusiewicz R, Zadeh L, Zurada J (eds) *Artificial intelligence and soft computing*, 1st edn. Springer, Berlin, Heidelberg, pp 213–220
- Chaira T (2011) A novel intuitionistic fuzzy *C* means clustering algorithm and its application to medical images. *Appl Soft Comput* 11(2):1711–1717
- Chaira T, Panwar A (2013) An Atanassov’s intuitionistic fuzzy kernel clustering for medical image segmentation. *Int J Comput Intell Syst* 17:1–11
- Cuong BC (2014) Picture fuzzy sets. *J Comput Sci Cybern* 30(4):409–420
- Davies DL, Bouldin DW (1979) A cluster separation measure. *IEEE Trans Pattern Anal Mach Intell* 2:224–227
- De Oliveira JV, Pedrycz W (2007) *Advances in fuzzy clustering and its applications*. Wiley, Chichester
- Graves D, Pedrycz W (2010) Kernel-based fuzzy clustering and fuzzy clustering: a comparative experimental study. *Fuzzy Sets Syst* 161(4):522–543
- Hung WL, Lee JS, Fuh CD (2004) Fuzzy clustering based on intuitionistic fuzzy relations. *Int J Uncertain Fuzziness Knowl-Based Syst* 12(4):513–529
- Hwang C, Rhee FCH (2007) Uncertain fuzzy clustering: interval type-2 fuzzy approach to *c*-means. *IEEE Trans Fuzzy Syst* 15(1):107–120
- Iakovidis DK, Pelekis N, Kotsifakos E, Kopanakis I (2008) Intuitionistic fuzzy clustering with applications in computer vision. *Lect Notes Comput Sci* 5259:764–774
- Ji Z, Xia Y, Sun Q, Cao G (2013) Interval-valued possibilistic fuzzy *C*-means clustering algorithm. *Fuzzy Sets Syst* 253:138–156
- Kaur P, Soni D, Gosain DA, India II (2012) Novel intuitionistic fuzzy *C*-means clustering for linearly and nonlinearly separable data. *WSEAS Trans Comput* 11(3):65–76
- Lin K (2014) A novel evolutionary kernel intuitionistic fuzzy *C*-means clustering algorithm. *IEEE Trans Fuzzy Syst* 22(5):1074–1087
- Linda O, Manic M (2012) General type-2 fuzzy *c*-means algorithm for uncertain fuzzy clustering. *IEEE Trans Fuzzy Syst* 20(5):883–897
- Mendel JM, John RB (2002) Type-2 fuzzy sets made simple. *IEEE Trans Fuzzy Syst* 10(2):117–127
- Son LH (2014a) Enhancing clustering quality of geo-demographic analysis using context fuzzy clustering type-2 and particle swarm optimization. *Appl Soft Comput* 22:566–584
- Son LH (2014b) HU-FCF: a hybrid user-based fuzzy collaborative filtering method in recommender systems. *Exp Syst Appl* 41(15):6861–6870
- Son LH (2014c) Optimizing municipal solid waste collection using chaotic particle swarm optimization in GIS based environments: a case study at Danang City, Vietnam. *Exp Syst Appl* 41(18):8062–8074
- Son LH (2015) DPFCM: a novel distributed picture fuzzy clustering method on picture fuzzy sets. *Exp Syst Appl* 42(1):51–66
- Son LH, Cuong BC, Lanzi PL, Thong NT (2012a) A novel intuitionistic fuzzy clustering method for geo-demographic analysis. *Exp Syst Appl* 39(10):9848–9859
- Son LH, Lanzi PL, Cuong BC, Hung HA (2012b) Data mining in GIS: a novel context-based fuzzy geographically weighted clustering algorithm. *Int J Mach Learn Comput* 2(3):235–238
- Son LH, Cuong BC, Long HV (2013) Spatial interaction-modification model and applications to geo-demographic analysis. *Knowl-Based Syst* 49:152–170
- Son LH, Linh ND, Long HV (2014) A lossless DEM compression for fast retrieval method using fuzzy clustering and MANFIS neural network. *Eng Appl Artif Intell* 29:33–42
- University of California (2007) UCI Repository of Machine Learning Databases. <http://archive.ics.uci.edu/ml>. Accessed 26 Nov 2014
- Vendramin L, Campello RJ, Hruschka ER (2010) Relative clustering validity criteria: a comparative overview. *Stat Anal Data Min* 3(4):209–235
- Xu Z (2012) Intuitionistic fuzzy clustering algorithms. In: Xu Z (ed) *Intuitionistic fuzzy aggregation and clustering*, 1st edn. Springer, Berlin, Heidelberg, pp 159–267
- Xu Z, Wu J (2010) Intuitionistic fuzzy *C*-means clustering algorithms. *J Syst Eng Electron* 21(4):580–590
- Zarandi MF, Gamasae R, Turksen IB (2012) A type-2 fuzzy *c*-regression clustering algorithm for Takagi–Sugeno system identification and its application in the steel industry. *Inf Sci* 187:179–203
- Zhao H, Xu Z, Wang Z (2013) Intuitionistic fuzzy clustering algorithm based on Boole matrix and association measure. *Int J Inf Technol Decis Mak* 12(1):95–118
- Zimmermann HJ (2001) *Fuzzy set theory-and its applications*. Springer, New York