Clustering Ensemble Based on Fuzzy Matrix Self-Enhancement

Xia Ji, Jiawei Sun, Jianhua Peng, Yue Pang and Peng Zhou

***Abstract*—Fuzzy clustering ensemble techniques have been proven to yield more accurate and robust clustering results, with the mainstream methods relying on the fuzzy co-association (FCA) matrix. However, the inherent issues of low-value density and uniform dispersion in the FCA matrix significantly affect the performance of fuzzy clustering ensembles, an aspect that has been overlooked. To address this issue, we propose a novel framework for fuzzy clustering ensemble based on fuzzy matrix self-enhancement (FMSE). Specifically, we initially employ singular value decomposition to extract the principal components of the FCA matrix, thereby alleviating its low-value density. Second, on the basis of the criterion of fuzzy entropy, we measure the fuzziness of samples, design a metric for the fuzzy representativeness of samples, and incorporate it into a fusion-weighted structure for the reconstruction of the FCA matrix, mitigating uniform dispersion. Subsequently, on the basis of the self-enhanced fuzzy matrix model, we utilize a prototype diffusion approach to identify core samples and gradually allocate remaining samples to obtain a consensus clustering solution. Extensive comparative experiments on benchmark datasets against state-of-the-art clustering ensemble methods demonstrate the effectiveness and superiority of the proposed approach. The codes of this paper are released in https://github.com/linshenwuqi/FMSE.**

***Index Terms*—Fuzzy clustering, clustering ensemble, Fuzzy co-association matrix.**

I. Introduction

CLUSTERING analysis is a pivotal and challenging technique in the fields of data mining and machine learning and aims to partition similar data objects into groups or clusters to unveil inherent patterns and structures within the data. Despite the widespread practical applications of numerous clustering algorithms [1]–[5], a solitary clustering method often falls short when confronted with complex and diverse datasets. Furthermore, these methods typically require raw data features as parameters for application. In contrast, clustering ensemble eliminates the need for direct access to raw data features. By combining multiple base clustering results, it can generate clustering outcomes that are more robust and accurate than those achieved by individual base clustering methods [6]. Consequently, clustering ensemble has become a focal point in unsupervised learning research and has proven effective across various clustering tasks, such as image segmentation [7], noise reduction [8], and time series analysis [9], making it a powerful way of enhancing clustering performance in scenarios where single clustering methods may encounter limitations.

Clustering ensemble research began with the introduction of CSPA, HGPA, and MCLA by Strehl and Ghosh in 2002 [10]. Many clustering ensemble methods have been proposed in recent years. These methods are based on ensemble strategies at the base clustering level and can be broadly categorized into three types: pairwise similarity-based methods [11]–[15], graph partitioning-based approaches [6], [16]–[19], and optimization-based methods [20]–[24]. With respect to the nature of clustering (crisp or fuzzy) at the base level, clustering ensemble is further divided into crisp clustering ensemble and fuzzy clustering ensemble. In crisp clustering ensemble, data objects are assigned to unique clusters within the base clustering. Fuzzy clustering ensemble allows data objects in the base clustering to be distributed among multiple clusters with varying membership degrees, where membership values range from [0, 1]. Crisp clustering ensemble may result in some information loss because of the inherent ambiguity in the cluster affiliation of data objects in real data distributions. Fuzzy clustering ensemble can handle inherent fuzziness and uncertainty in the data better than crisp clustering ensemble can. In clustering ensemble, the quality and diversity of base clustering are important factors for enhancing the final clustering quality [25]–[27]. Fuzzy base clustering efficiently preserves diverse information, which is a crucial element for the overall improvement of ensemble performance. Hence, introducing fuzzy clustering as a base clustering component in clustering ensemble is appropriate.

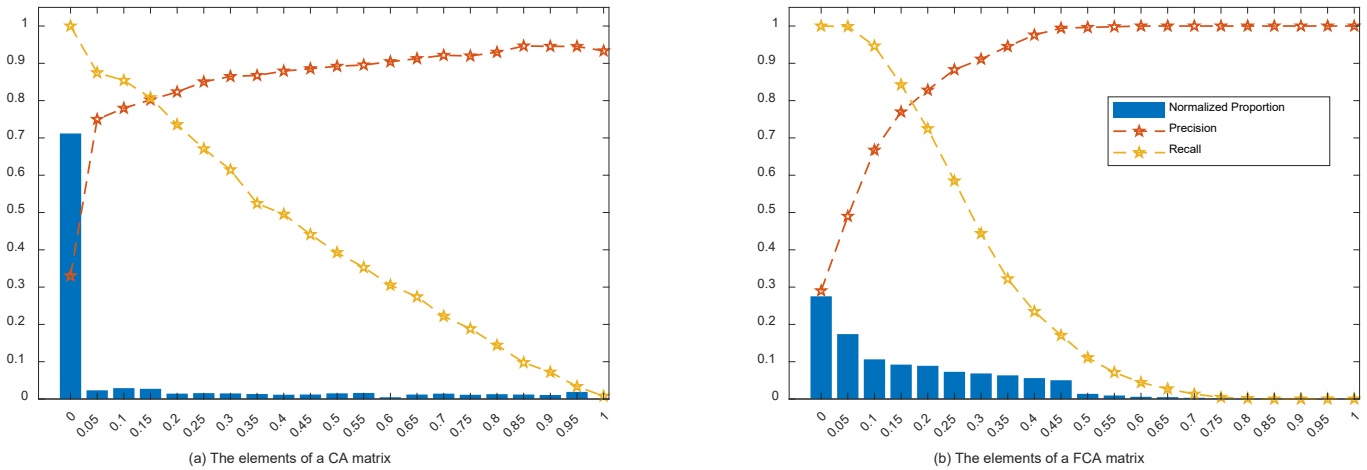
In 2008, Punera and Ghosh [10] introduced sCSPA, sHBGF, and sMCLA, which initiated the research on fuzzy clustering ensemble. Following this, numerous fuzzy clustering ensemble algorithms have been proposed. In [28], Su et al. proposed three link-based pairwise fuzzy similarity matrices and applied hierarchical clustering algorithms with complete linkage and average linkage on these matrices to derive consensus clustering results. In [15], Li et al. proposed a self co-association prototype propagation (SCPP) algorithm. This algorithm utilizes the diagonal elements of the FCA matrix to evaluate the local density of samples, identifying a set of prototype samples, and completing clustering by assigning the remaining samples based on sample similarity.

A thorough examination of current research in fuzzy clustering ensemble indicates that the majority of existing methods draw inspiration from the concept of the co-association (CA) matrix commonly employed in crisp clustering ensemble. These methods first construct an FCA matrix and then ap-

Xia Ji, Jiawei Sun, Jianhua Peng and Peng Zhou are with School of Computer Science and Technology, Anhui University, Hefei, 230601, China; Xia Ji and Peng Zhou are also with the Anhui Provincial International Joint Research Center for Advanced Technology in Medical Imaging. (Email: jixia@ahu.edu.cn and jwsun@stu.ahu.edu.cn)

Yue Pang is with School of Computer and Information Engineering, Henan University, Kaifeng, 450046, China.

(Corresponding author: Peng Zhou.) Email: zhoupeng@ahu.edu.cn



(a) The elements of a CA matrix (b) The elements of a FCA matrix

Fig. 1. Visualization of CA and FCA matrices on seeds dataset.

ply graph partitioning or hierarchical clustering methods on the basis of the FCA matrix to derive the final clustering results. The FCA matrix records the frequency with which multiple base clustering results assign two samples to the same partition, enabling it to quantify the fuzzy relationships between samples. Therefore, the FCA matrix can be regarded as the adjacency matrix of a similarity graph or matrix for data objects. However, the FCA matrix itself exhibits two characteristics that can impact the final performance of fuzzy clustering ensemble.

- **Low-value density:**
  Typically, higher CA values provide more reliable information about the fuzzy relationship between two samples, whereas lower CA values may introduce unreliable information or even noise. However, the FCA matrix contains numerous low-value elements. Unlike in the CA matrix, where zero values provide exact information, these low-value elements in the FCA matrix may contain noise and unreliable information, significantly impacting subsequent clustering outcomes.
- **Uniform dispersion:**
  The range of CA values is [0, 1], where element values of 0 or 1 indicate a deterministic relationship between samples. As the element value approaches 0.5, the fuzziness between samples increases. In contrast to the high sparsity of the CA matrix, the FCA matrix has a relatively uniform distribution of CA values. This finding implies that most samples have high fuzziness, with their membership spread evenly among various fuzzy clusters. The resulting partition lacks a clear clustering tendency, which is detrimental to the overall ensemble performance.

As illustrated in Fig. 1, we employed the K-means and fuzzy C-means (FCM) algorithms on the real dataset ”seeds” to generate 100 crisp clustering results and 100 fuzzy clustering results separately. Subsequently, we constructed the corresponding CA matrix and the FCA matrix. The bar charts in Fig. 1 depict the distribution of element values in the CA and FCA matrices on the ”seeds” dataset. The line charts illustrate variations in precision and recall corresponding to different value ranges in the CA and FCA matrices. The horizontal and vertical axes represent elements in the matrix within different ranges and their respective proportions. The graph shows that higher CA values are correlated with higher precision and lower recall, and vice versa. This finding suggests that, in both the CA matrix and the FCA matrix, lower CA values contain unreliable information. Furthermore, a comparison of Figs. 1(a) and 1(b) shows that, unlike the CA matrix constructed from crisp clustering results, which predominantly contains zero elements (high sparsity), the FCA matrix has uniform dispersion with numerous low-value elements. This finding indicates that the membership distribution of most samples among various fuzzy clusters is relatively even, signifying higher fuzziness. This phenomenon results in the emergence of fuzzy cluster boundaries in the final consensus results, making it difficult to differentiate between different clusters.

In this paper, we propose a clustering ensemble algorithm based on fuzzy matrix self-enhancement (FMSE), aiming to mitigate the adverse impact of the two characteristics of the FCA matrix on clustering ensemble performance. Specifically, to address the issue of low-value density in the FCA matrix, we incorporate singular value decomposition (SVD) to extract the principal components of the FCA matrix. SVD decomposes the FCA matrix into singular values and their corresponding left and right singular vectors. Singular values signify the importance or variability of the data, while the left and right singular vectors represent the distribution of data in the feature space. By selecting components with relatively large singular values and their corresponding singular vectors, we can extract the main features of the FCA matrix. Filtering out components corresponding to small singular values, which are indicative of noise data, facilitates denoising and redundancy removal in the FCA matrix. To alleviate the issue of uniform dispersion in the FCA matrix, we design the fuzzy representativeness indicator (FRI) of the samples based on fuzzy entropy and reconstruct the FCA matrix through representativeness weighting. Representative weighting enables us to more effectively leverage the uncertainty information in the fuzzy partition matrix, giving emphasis to more representative samples and thereby reducing the impact of unreliable data. By introducing SVD and representative weighting to enhance the FCA matrix,

we define a novel fuzzy matrix model called the self-enhanced FCA (EFCA) matrix, which aims to effectively describe the fuzzy relationship of samples. Subsequently, on the basis of the EFCA matrix model, we propose a fuzzy clustering ensemble algorithm based on self co-association prototype diffusion. This algorithm uses the EFCA matrix to gauge the local density of samples, applies prototype clustering principles to identify core samples, and allocates remaining samples based on the fuzzy relationship of core samples to obtain consensus clustering results. Extensive comparative experiments with 12 benchmark datasets against 12 state-of-the-art clustering ensemble algorithms, including four crisp clustering ensemble and eight fuzzy clustering ensemble algorithms, demonstrated the superiority of our proposed method.

For clarity, the main contributions of this paper can be summarized as follows:

1) Addressing the issues of low-density values and uniform dispersion in the FCA matrix, we introduce SVD to extract the principal components of the FCA matrix and design a fuzzy entropy-driven indicator of sample representativeness. Consequently, we propose a novel self-enhanced fuzzy matrix model called EFCA, which reasonably characterizes the fuzzy correlations among samples, thus capturing the essential features of the original data effectively.
2) On the basis of the EFCA matrix model, we present a clustering ensemble algorithm named fuzzy matrix self-enhancement (FMSE). This algorithm utilizes the self co-association of samples to measure local density for discovering prototype samples. It then employs the information from the EFCA matrix to gradually diffuse marginal samples to the prototype sample set, achieving efficient clustering for complex data structures.
3) Extensive comparative experiments conducted on multiple benchmark datasets against state-of-the-art clustering ensemble algorithms demonstrate the effectiveness, superiority, and robustness of FMSE.

This paper is organized as follows: Section II presents the formulaic background knowledge of the proposed methodology. Section III describes the proposed fuzzy clustering ensemble method in detail. Section IV demonstrates the effectiveness and superiority of the proposed method through an analysis of experimental results. Section V summarizes the key findings of this paper.

II. Preliminaries and Related Work

In this section, we first introduce the symbols used in this paper and relevant foundational theories. Subsequently, we describe the fuzzy clustering ensemble problem and review some representative works. The symbols used in this paper are presented in Table I.

TABLE I
SUMMARY OF NOTATIONS

Notation	Description
X	Dataset
x_i	ith data object in dataset X
N	Number of data objects in dataset X
D	Number of features of data object x_i
M	Number of fuzzy base clustering in fuzzy set Π_f
Π_c , Π_f	Sets of M crisp and fuzzy clustering results
H^m, F^m	mth base clustering result
C	Number of clusters in fuzzy set Π_f
C_i^m	ith cluster in the mth base clustering
$n^{(m)}$	Number of clusters in the mth base clustering
CA, FCA	(Fuzzy) co-association matrix
U, V	Left and right singular vector matrices
S	Singular value matrix
σ	Singular value
energy	Information content in S
k	Number of retained principal components
α	Parameter that adjusts the principal components
θ	Parameter that adjusts the fuzziness impact on FRI
EFCA	Self-enhanced FCA matrix model
$\rho(x_i)$	Local density of data object x_i
δ_i	Representativeness of data object x_i
$\mathbb{R}$	Prototype sample set (i.e., set of initial clusters)
R_s	sth initial cluster in the prototype sample set
F^*	The final clustering result

A. Singular Value Decomposition

SVD [29] is a commonly used technique for matrix decomposition and dimensionality reduction and has widespread applications in fields such as data analysis, signal processing, image compression, and recommendation systems. SVD decomposes a matrix into the product of three matrices, each containing essential information from the original matrix, making it useful for feature extraction and data compression.

Definition 1. For a matrix A of size $m \times n$, its singular value decomposition can be expressed as:

$$A = U\Sigma V^T \tag{1}$$

where U is an $m \times m$ orthogonal matrix, satisfying $U \cdot U^T = U^T \cdot U = I_m$, I_m is the m-dimensional identity matrix.

Σ is an $m \times n$ diagonal matrix, with non-negative and decreasing elements on the diagonal, satisfying $\sigma_1 \geq \ldots \geq \sigma_r > 0$, where r is the rank of A, i.e., the number of singular values. V is an $n \times n$ orthogonal matrix, satisfying $V \cdot V^T = V^T \cdot V = I_m$. The product $U\Sigma V^T$ is the SVD of matrix A, σ_i is the singular value of matrix A, and the column vectors of U and V are the left singular vectors and right singular vectors, respectively, corresponding to the singular values in Σ.

Singular values in SVD represent the importance or variability of data, while left and right singular vectors depict the distribution of data in the feature space. The main features of the FCA matrix can be extracted by selecting components with larger singular values and their corresponding singular vectors. Filtering components corresponding to smaller singular values, which are associated with noise data, can achieve denoising and eliminate redundancy.

B. Fuzzy Entropy

Fuzzy entropy is a metric used to quantify the uncertainty or degree of disorder in fuzzy sets, extending the concept of Shannon entropy [30] to fuzzy sets. It has wide applications in areas such as fuzzy logic, fuzzy clustering, and fuzzy decision-

making, providing valuable insights into the uncertainty associated with fuzzy data for use in fuzzy inference and decision-making.

Definition 2. For a fuzzy set A, the fuzzy entropy $H(A)$ is defined as:

$$H(A) = -\sum_{x \in A} \mu_A(x) \log_2 \mu_A(x) \tag{2}$$

where $\mu_A(x)$ is the membership function of the element x in the fuzzy set, and it satisfy the following two constraints:

- Regularity: $\sum_{x \in A} \mu_A(x) = 1$
- Non-negativity: $0 \leq \mu_A(x) \leq 1$

Fuzzy entropy measures the uncertainty of elements in a fuzzy set. Higher fuzzy entropy indicates a more dispersed or uncertain distribution of membership degrees within the fuzzy set. Conversely, lower fuzzy entropy suggests a more concentrated or certain distribution of membership degrees among the elements of the fuzzy set. The distinct contributions of different samples to the clustering ensemble can be reflected better by calculating the fuzzy entropy based on membership degrees for each sample and incorporating it into a weighted structure as a representative indicator. This approach emphasizes more representative samples, reduces the impact of unreliable data, and thereby improves the quality and reliability of the FCA matrix.

C. Fuzzy Clustering Ensemble

Let $X = \{x_1, x_2, \ldots, x_N\}$ denote a dataset containing N data objects, where x_i represents the ith data object, each consisting of D features. The fuzzy partition result of a dataset can be represented as a 2D matrix F of size $N \times K$, where N is the number of data objects and K is the number of clusters of partitioning. Matrix $F(X)$ satisfies the following conditions:

$$\begin{aligned} &\forall i \forall j, i \in 1, \ldots, N, j \in 1, \ldots, K: \\ &\quad F_j(x_i) \in [0,1] \quad \text{and} \quad \sum_{j=1}^{K} F(x_i) = 1 \end{aligned} \tag{3}$$

A set of fuzzy clustering sets consisting of M fuzzy partitions can be denoted as follows:

$$\begin{aligned} \Pi_f &= \left\{F^1, \ldots, F^M\right\} \\ &= \{F^1_{*,1}, F^1_{*,2}, \ldots, F^1_{*,n^1}, \ldots, F^M_{*,1}, \ldots, F^M_{*,n^m}\} \end{aligned} \tag{4}$$

where the fuzzy partition $F^m = \{C^m_1, \ldots, C^m_{n^m}\}$ denotes the $n^{(m)}$ fuzzy clusters in the mth fuzzy partition. For each $x_i \in X$, $F^m_j(x_i)$ denotes the probability (known as membership) that data object x_i in fuzzy partition F^m belongs to fuzzy cluster C^m_j.

The research objective of fuzzy clustering ensemble is to derive an ensemble clustering solution F^* based on a set of fuzzy clustering results Π_f. This process primarily involves two crucial steps: generating multiple distinct fuzzy clustering and integrating the results of these fuzzy clustering to produce the final partition. During the generation phase, the base clustering members should exhibit diversity so that the effectiveness of the ensemble is enhanced, and each base clustering should have a certain level of accuracy [31]. Generally, in fuzzy clustering ensembles, base clustering is typically obtained using the FCM algorithm by randomly selecting different numbers of clusters and varying fuzzy exponents. In what follows, we introduce some representative methods.

In [32], Popescu et al. introduced random projection fuzzy C-means clustering (RPFCM), obtaining final clustering results by applying the FCM to multiple fuzzy partition matrices. Avogadri et al. [33] applied FCM to the rows of the CA matrix to obtain ensemble clustering results, which is a method known as EFCM. In [34], Ye et al. presented an enhanced version of EFCM, EFCM-S, which applies spectral clustering to a set of fuzzy clustering collections to derive final clustering results. In [35], Dhiloon et al. introduced the fuzzy clustering algorithm ITK, treating each row of fuzzy base clustering sets as features in a new space. They calculated the distance between data objects using KL divergence [36] and then applied a K-means-like algorithm on the distance matrix to derive the final ensemble clustering results. In [37], Zhou et al. proposed an active clustering ensemble method based on self-paced learning, which optimizes the ensemble model by selecting and labeling unreliable data to improve clustering performance. In [38], Mojarad et al. introduced a novel fuzzy cluster similarity measure, applying hierarchical clustering on the fuzzy cluster similarity matrix to obtain clustering results.

In [39], Bedalli et al. proposed a heterogeneous fuzzy clustering ensemble approach to increase the robustness of fuzzy clustering results by applying several fuzzy clustering algorithms to generate the base clustering and then applying an FCM on the FCA matrix to obtain the final clusterings. Berikov et al. [40] proposed a probability model-based fuzzy clustering ensemble method, utilizing hierarchical clustering on an FCA matrix to obtain the final ensemble clustering results. Bagherinia et al. [25] clustered a set of fuzzy clustering results by using FCM. They proposed a method that simultaneously considers the quality and diversity of fuzzy clusters to select a subset of fuzzy clustering results. Then, they applied hierarchical clustering to the FCA matrix generated from the fuzzy clustering subset to obtain the final clustering results. In [41], on the basis of the distinct contributions of different fuzzy clusters to the ensemble, Bagherinia et al. introduced a reliability-driven metric for measuring fuzzy cluster quality called RDCI, which was incorporated into a weighted structure to obtain a weighted FCA matrix. Finally, clustering results were generated based on hierarchical clustering and bipartite graph partitioning, which are two consensus functions.

Multiview clustering can be treated as a special ensemble clustering using different views. Many multi-view clustering methods have been proposed to integrate multiple fuzzy partitions of multiple views. In [42], Zhang et al. developed a novel multiview and multiexemplar fuzzy clustering approach (M2FC), and demonstrated that M2FC has a theoretical guarantee of enhanced clustering performance. In [43], Kang et al. utilized anchors and bipartite graphs to establish the relationship between samples and anchors, proposing a structured graph learning framework for scalable subspace clustering. In [44], Lin et al. introduced a novel multi-view attribute graph clustering (MAGC) framework, which efficiently clusters multi-view attribute graph data into distinct groups. In [45], the author proposed four multiview fuzzy

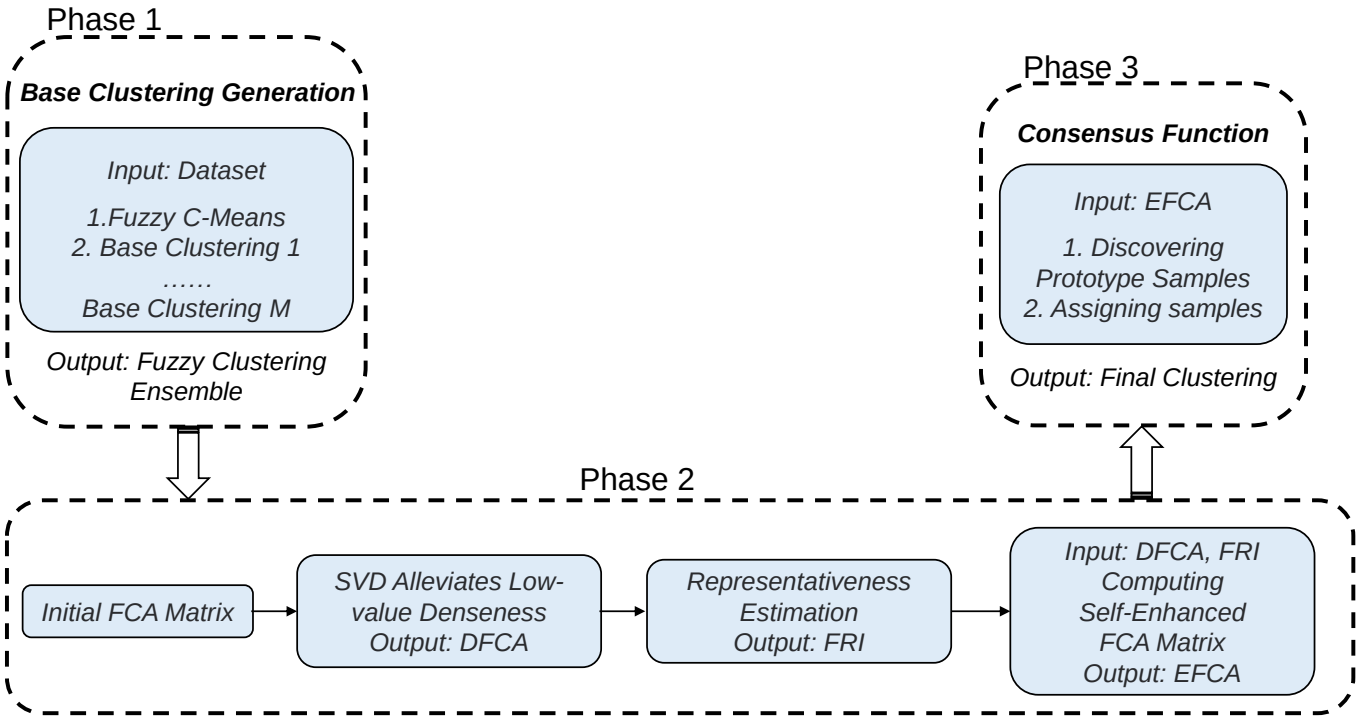


Fig. 2. The illustration of the proposed FMSE.

clustering frameworks for clustering single-view data with missing values.

In recent years, the CA matrix has played a crucial role in similarity matrix-based ensemble methods. Through the analysis of the CA matrix, a deeper understanding of the relationships among data samples and their consistency across multiple clustering results can be obtained. However, due to the fact that fuzzy clustering results differ from crisp clustering's label vectors, they can only be expressed as membership matrices based on clusters. Therefore, we cannot obtain the FCA matrix by simply recording the occurrences of a pair of data objects within the same fuzzy cluster.

Definition 3. Given an ensemble Π_f with M fuzzy base clustering, the FCA matrix can be represented as follows:

$$\text{FCA} = [fca_{ij}]_{N\times N} \tag{5}$$

$$fca_{ij} = \frac{1}{M}\sum_{m=1}^{M}\sum_{c=1}^{n^{(m)}} F_{i,c}^{m}\cdot F_{j,c}^{m} \tag{6}$$

Typically, the calculation expression for the FCA matrix is given by $\text{FCA} = \frac{1}{M}\Pi_f \times \Pi_f^T$, where Π_f^T is the transpose of Π_f.

Despite numerous clustering ensemble algorithms based on the CA matrix, research on fuzzy clustering ensemble based on the FCA matrix is still in its infancy. Moreover, the performance of fuzzy clustering ensemble needs to be improved because of the low-density and uniformly dispersed characteristics of the FCA matrix.

III. Proposed approach

This paper proposes a novel fuzzy clustering ensemble algorithm based on self-enhanced fuzzy matrix, as illustrated in the algorithm framework shown in Fig. 2.

The entire algorithm can be divided into three stages. In Stage 1, a fuzzy clustering ensemble is generated by employing FCM clustering algorithms with different initializations, and this ensemble serves as the input for the second stage. Stage 2 constitutes the core of the algorithm, encompassing four steps. First, an FCA matrix is constructed. Subsequently, the obtained FCA matrix undergoes SVD to extract primary features and reduce noise, as detailed in Section III-A. Then, the representativeness of samples regarding the ensemble as a whole is calculated, as indicated in Section III-B. Finally, the EFCA matrix is constructed by weighting based on sample representativeness, as discussed in Section III-C. In Stage 3, the local density of samples is measured by the diagonal elements of the EFCA matrix. The core samples are discovered using density peak clustering. Subsequently, the remaining samples are assigned based on the information from the EFCA matrix, gradually diffusing into the prototype sample set to obtain the final clustering result, as described in Section III-D.

A. SVD Alleviates Low-value Denseness

After obtaining the FCA matrix, we introduce SVD to denoise the FCA matrix and extract key features, resulting in the DFCA matrix, to alleviate the low-value density of the FCA matrix.

We start by performing SVD on the initial FCA matrix as follows:

$$\text{FCA} = U\cdot S\cdot V^T \tag{7}$$

where U and V are the left and right singular vector matrices, respectively, and S is the singular value matrix with diagonal elements σ, arranged in descending order $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_N$.

Next, we need to calculate the information content carried by different singular values in the matrix S

$$\sigma = \text{diag}(S) \tag{8}$$

$$\text{energy} = \frac{\text{cumsum}(\sigma_i^2)}{\sum_{i=1}^{N} \sigma_i^2} \tag{9}$$

where $\text{cumsum}\left(\sigma_i^2\right)$ represents the cumulative sum of squared singular values, i.e.,

$$\text{cumsum}(\sigma_i^2) = \left[\sigma_1^2, \left(\sigma_1^2 + \sigma_2^2\right), \ldots, \left(\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_N^2\right)\right] \tag{10}$$

Then, on the basis of parameter α, we determine the number k of principal components to retain

$$k = \text{argmin}\left(\text{energy} \geq \alpha\right) \tag{11}$$

Finally, the matrices obtained from the SVD decomposition are truncated to obtain the DFCA matrix, preserving the first k principal components

$$\text{DFCA} = U_k \cdot S_k \cdot V_k^T \tag{12}$$

where U_k and V_k are the first k columns of matrices U and V, and S_k is the diagonal matrix retaining the first k singular values.

B. Fuzzy Representativeness Estimation

To mitigate the impact of uniform dispersion, we utilize fuzzy entropy to calculate the fuzziness of each sample. Consequently, we estimate the representativeness of the samples and incorporate them into the weighted structure. Specifically, we first calculate the fuzziness of each sample with respect to the fuzzy clusters.

Definition 4. Given a data object x_i in the set and a fuzzy cluster C_j^m(where $C_j^m \in F^m$), the fuzziness of data object x_i relative to fuzzy cluster C_j^m is calculated as:

$$FE(x_i, C_j^m) = -\sum_{x_i \in X} F_j^m(x_i) \log_2 F_j^m(x_i) \tag{13}$$

where, $F_j^m(x_i)$ represents the membership degree of data object x_i to the jth fuzzy cluster in the mth base partition.

In this context, we assume that the base clustering and clusters in the fuzzy set are independent [46]. Therefore, the fuzziness of data object x_i relative to the entire set can be obtained by summing the fuzziness of x_i relative to fuzzy cluster C_j^m.

Definition 5. Given a fuzzy ensemble Π_f with M base clustering, the fuzziness of data object x_i relative to the entire ensemble Π_f is defined as

$$FE(x_i, \Pi_f) = \sum_{m=1}^{M} \sum_{j=1}^{n^m} FE(x_i, C_j^m) \tag{14}$$

where, $n^{(m)}$ represents the number of fuzzy clusters in the mth base clustering result.

After obtaining the fuzziness of each sample in the fuzzy ensemble, we further introduce the concept of FRI, which measures the representativeness of samples by considering the fuzziness of samples relative to the ensemble.

Definition 6. Given a fuzzy ensemble Π_f with M base clustering, the fuzzy representativeness of data object x_i relative to the entire ensemble Π_f (i.e., FRI) is defined as

$$FRI(x_i, \Pi_f) = e^{-\frac{FE(x_i, \Pi_f)}{(M-1)\cdot\theta}} \tag{15}$$

The parameter $\theta > 0$ is used to adjust the influence of sample fuzziness (FE) on the FRI. Section IV-D provides a more detailed discussion on the values of the parameter θ on the datasets used in this paper.

According to the above definition, because $FE^{\Pi}(x_i) \in [0, +\infty)$ for any $x_i \in X$, $FRI(x_i) \in (0, 1]$. Clearly, a small fuzziness of data object x_i indicates that its FRI value is larger. When the fuzziness of data object x_i reaches its minimum value, i.e., $FE^{\Pi}(x_i) = 0$, its FRI will reach its maximum value, i.e., $FRI(x_i) = 1$. As the fuzziness of a sample approaches infinity, its FRI approaches zero.

C. Self-Enhanced Fuzzy Matrix

On the basis of the DFCA matrix from Equation (12) and the FRI obtained in Section III-B, we can compute the EFCA.

Definition 7. Given a set Π_f containing M fuzzy clustering results, the self-enhanced FCA matrix, based on sample representativeness weighting, is defined as

$$\text{EFCA} = \{efca_{ij}\}_{N \times N} \tag{16}$$

$$efca_{ij} = \frac{FRI_i \cdot DFCA_{ij} + FRI_j \cdot DFCA_{ji}}{2} \tag{17}$$

where FRI_i represents the representativeness of data object x_i, obtained through Equation (15), and $DFCA_{ij}$ represents the FCA value between the ith and jth data objects in the matrix after extracting the main features through SVD, obtained through Equation (12).

The introduction of the FRI metric serves as a weighting factor for sample allocation after denoising and extracting main components from the FCA matrix. Intuitively, data objects with lower fuzziness (higher FRI values) contribute to a more stable data structure. Through the fusion of a weighted strategy, the EFCA matrix not only alleviates the low-value density in the FCA matrix but also effectively addresses the issue of uniform distribution, thereby enhancing the quality of ensemble clustering.

D. Consensus Functions

After obtaining the self-enhanced EFCA matrix model, we measure the local density of samples using the diagonal elements of the EFCA matrix. Subsequently, we identify a set of stable core samples: those with high local density and a considerable distance from samples with higher local density. On the basis of the fuzzy relationships among samples described in the EFCA matrix, we allocate the remaining samples to complete the ensemble clustering task. We first measure the local density ρ_{x_i} of sample x_i based on the self co-association value in the EFCA matrix

$$\rho_{xi} = efca_{ii} \tag{18}$$

where $efca_{ii}$ represents the diagonal element of the EFCA matrix.

Next, we calculate the representativeness δ of samples with local density ρ greater than its average value $\bar{\rho} = \frac{1}{N}\sum_{i=1}^{N}\rho_i$

$$\delta_i = \min_{j:\rho_j > \rho_i} (1 - efca_{ij}) \tag{19}$$

By sorting samples from high to low representativeness, we obtain the prototype sample set (i.e., the initial cluster set)

$$\mathbb{R} = \{R_1, R_2, \ldots, R_k | 1 \le k \le N\} \tag{20}$$

Algorithm 1 Clustering Ensemble based on Fuzzy Matrix Self-Enhancement (FMSE)

INPUT: Fuzzy co-association matrix **FCA**, Final cluster number **K**, Thresholds θ, α.

OUTPUT: Final clustering result $\mathbf{F}^*$

Process:
Extract principal components of the FCA matrix according to Equation (12).
Compute the fuzziness of samples to the entire set using Equation (14).
Calculate the FRI of each sample in the set based on Equation (15).
Compute the EFCA matrix using Equation (17).
Discover the set of prototype samples according to Equation (20).
Assign the remaining samples using Equation (25).
Merge redundant clusters until the cluster number equals **K** to obtain the ensemble clustering result $\mathbf{F}^*$.

Finally, we allocate the remaining samples to the above prototype sample set, causing the prototype sample set to gradually diffuse. Before doing so, we need to calculate the distance between the remaining samples and the prototype sample set and assign each sample to the nearest set until all samples are allocated. The proximity of sample x_i to the prototype sample set (R_s) is defined as

$$ps(x_i, \mathbb{R}) = \max_{b=1,2,\ldots,k} sim(x_i, R_s) \tag{21}$$

where $sim(x_i, R_s)$ is the fuzzy similarity between sample x_i and prototype samples in the initial cluster R_s. In this paper, we propose two fuzzy similarity measures, namely, average based and voting based, which are defined as follows:

Average-based fuzzy similarity measure:

$$sim_a(x_i, R_s) = \max_{x_b \in R_s} efca_{ib} \tag{22}$$

Voting-based fuzzy similarity measure:

$$sim_v(x_i, R_s) = \sum_{m=1}^{M} \mathbb{I}_v(x_i, R_s, F^m) \tag{23}$$

Where:

$$\mathbb{I}_v(x_i, R_s, F^m) = \begin{cases} 1, & \max_{h=1,2,\ldots,k}\{v(x_i, R_s, F^m)\} = v(x_i, R_s, F^m) \\ 0, & \text{otherwise} \end{cases} \tag{24}$$

$$v(x_i, R_s, F^m) = \max_{x_b \in R_s} \sum_{c=1}^{n^m} f_{i,c}^b \times f_{j,c}^b \tag{25}$$

With the assumption that x_i is an unallocated sample, the allocation process involves finding the cluster to which it should be assigned, which is calculated as

$$j = \arg\max(ps(x_i, \mathbb{R})) \tag{26}$$

Then, the sample is assigned to the nearest cluster

$$R_s = R_s \cup x_i \tag{27}$$

If the number of prototype sample sets after allocation exceeds the required final number of clusters, then hierarchical clustering is used with a single linkage measure to merge prototype sample sets until the desired number of clusters is achieved.

On the basis of the two similarity metrics sim_a and sim_v mentioned above, this paper proposes two fuzzy clustering ensemble algorithms—FMSE-a and FMSE-v—to accomplish the final fuzzy clustering integration process. The entire algorithmic procedure is described in Algorithm 1.

IV. Experiments

In this section, we conducted a series of comparative experiments with 12 state-of-the-art clustering ensemble algorithms, comprising eight fuzzy clustering ensemble algorithms and four crisp clustering ensemble algorithms, as detailed in Section IV-B. Additionally, we examined the influence of different ensemble sizes on algorithm performance and conducted a sensitivity analysis of our method's parameters, as outlined in Sections IV-D and IV-E. Furthermore, in Sections IV-F and IV-G, we explored the time performance of the algorithms and discussed the necessity of the two proposed strategies through an ablation study, respectively.

Due to space constraints, we present only the ACC and NMI metric comparisons in the main manuscript. Please refer to the supplementary materials for the ARI metric experimental results. Additionally, the supplementary file includes an ablation study of the FMSE-v algorithm across all three metrics.

TABLE II
DESCRIPTION OF THE BENCHMARK DATA SETS

Number	Datasets	#data-objects	#features	#classes
1	Seeds	210	7	3
2	Heart	270	13	2
3	Ecoli	336	7	8
4	Climate	540	18	2
5	Diabetes	768	8	2
6	WDBC	569	30	2
7	CMC	1473	9	3
8	BCW	683	9	2
9	Waveform	5000	21	3
10	Spambase	4601	57	2
11	SMK_CAN_187	187	19993	2
12	LetterRecognition	20000	16	26

A. Experimental Set-up

All experiments were conducted using MATLAB R2021b on a Windows system with a 2.1 GHz CPU and 24GB of memory.

Data sets: we utilized 12 benchmark datasets from the UCI Machine Learning Repository [1] and the Bioinformatics Feature

[1] https://archive.ics.uci.edu/

TABLE III
THE ACC RESULTED FROM DIFFERENT ALGORITHMS, PRESENTED AS "MEAN ± STANDARD DEVIATION"

Method	Seeds	Heart	Ecoli	Climate	Diabetes	WDBC
HGPA [*MLR*, 2002]	0.8848 ± 0.0015	0.6107 ± 0.0129	0.4911 ± 0.0460	0.5159 ± 0.0134	0.5089 ± 0.0077	0.6469 ± 0.0917
PTA-CL [*TKDE*, 2016]	0.6967 ± 0.0962	0.6000 ± 0.0356	0.5098 ± 0.0464	0.6472 ± 0.1021	0.5966 ± 0.0490	0.7459 ± 0.0669
CESC [*CC*, 2021]	0.8171 ± 0.1065	0.5881 ± 0.0405	0.6503 ± 0.0982	0.6293 ± 0.0931	0.5753 ± 0.0303	0.8480 ± 0.0263
CMSE [*TNNLS*, 2023]	0.7490 ± 0.1283	0.5989 ± 0.0382	$\mathbf{0.7693 \pm 0.0446}$	0.5909 ± 0.0899	0.5299 ± 0.0453	0.7431 ± 0.0936
HC-a [*FUZZ-IEEE*, 2014]	0.6819 ± 0.0118	0.5519 ± 0.0252	0.4161 ± 0.0554	0.5983 ± 0.1017	0.5801 ± 0.0374	0.6986 ± 0.0767
HC-c [*FUZZ-IEEE*, 2014]	0.6943 ± 0.0228	0.5611 ± 0.0215	0.4268 ± 0.0524	0.5835 ± 0.0800	0.5751 ± 0.0305	0.6921 ± 0.0675
RPFCM [*FUZZ-IEEE*, 2015]	0.8705 ± 0.0220	0.5767 ± 0.0258	0.6312 ± 0.0232	0.5278 ± 0.0212	0.5147 ± 0.0161	0.8363 ± 0.0045
FCSPA [*AAI*, 2008]	0.8800 ± 0.0100	0.6095 ± 0.0037	0.4917 ± 0.0061	0.5148 ± 0.0000	0.5132 ± 0.0018	0.8274 ± 0.0186
RBEACE_AL [*FSS*, 2021]	0.8024 ± 0.0471	0.5948 ± 0.0346	0.6402 ± 0.0627	0.8749 ± 0.0143	0.5711 ± 0.0535	0.7286 ± 0.1511
RBEACE_CL [*FSS*, 2021]	0.7886 ± 0.0496	0.5963 ± 0.0335	0.5458 ± 0.0550	0.6963 ± 0.1171	0.5322 ± 0.0292	0.7351 ± 0.1314
SCPP-a [*TFS*, 2023]	0.7411 ± 0.0506	0.5907 ± 0.0382	0.6804 ± 0.0537	0.5941 ± 0.0785	0.6026 ± 0.0676	0.7961 ± 0.0780
SCPP-v [*TFS*, 2023]	0.7376 ± 0.0626	0.5822 ± 0.0333	0.6890 ± 0.0560	0.7513 ± 0.1192	0.6012 ± 0.0678	0.7970 ± 0.0802
FMSE-a	$\mathbf{0.8875 \pm 0.0978}$	$\mathbf{0.6220 \pm 0.0266}$	0.7116 ± 0.0147	$\mathbf{0.8815 \pm 0.0049}$	$\mathbf{0.6206 \pm 0.0582}$	$\mathbf{0.8487 \pm 0.0296}$
FMSE-v	$\mathbf{0.8862 \pm 0.0294}$	$\mathbf{0.6220 \pm 0.0234}$	$\mathbf{0.7134 \pm 0.0202}$	$\mathbf{0.8837 \pm 0.0050}$	$\mathbf{0.6135 \pm 0.0600}$	$\mathbf{0.8508 \pm 0.0389}$

Method	CMC	BCW	Waveform	Spambase	SMK	LR
HGPA [*MLR*, 2002]	0.3876 ± 0.0093	0.7204 ± 0.0708	0.5265 ± 0.0528	0.5236 ± 0.0195	0.6198 ± 0.0181	0.2287 ± 0.0159
PTA-CL [*TKDE*, 2016]	0.3859 ± 0.0092	0.6998 ± 0.0565	0.5844 ± 0.0585	0.6684 ± 0.0333	0.5390 ± 0.0455	$\mathbf{0.2966 \pm 0.0000}$
CESC [*CC*, 2021]	0.4052 ± 0.0225	0.8002 ± 0.0546	0.6109 ± 0.1089	0.6104 ± 0.0107	0.5417 ± 0.0435	0.2208 ± 0.0000
CMSE [*TNNLS*, 2023]	0.4206 ± 0.0155	0.7346 ± 0.0837	$\mathbf{0.6822 \pm 0.0922}$	0.6073 ± 0.0000	0.5599 ± 0.0434	N/A
HC-a [*FUZZ-IEEE*, 2014]	0.3847 ± 0.0144	0.6624 ± 0.0887	0.4751 ± 0.0267	0.5979 ± 0.0570	0.5727 ± 0.0334	0.1243 ± 0.0096
HC-c [*FUZZ-IEEE*, 2014]	0.3921 ± 0.0148	0.6951 ± 0.0864	0.4411 ± 0.0247	0.5721 ± 0.0395	0.5877 ± 0.0367	0.1728 ± 0.0178
RPFCM [*FUZZ-IEEE*, 2015]	0.3909 ± 0.0066	$\mathbf{0.8901 \pm 0.0085}$	0.6100 ± 0.0101	0.6744 ± 0.0024	0.5995 ± 0.0240	0.0760 ± 0.0200
FCSPA [*AAI*, 2008]	0.3860 ± 0.0024	0.8253 ± 0.0072	0.3417 ± 0.0028	$\mathbf{0.6890 \pm 0.0007}$	0.6070 ± 0.0058	0.1765 ± 0.0285
RBEACE_AL [*FSS*, 2021]	0.3833 ± 0.0042	0.7671 ± 0.1065	0.5072 ± 0.0170	0.6175 ± 0.0479	0.5171 ± 0.0062	0.0587 ± 0.0024
RBEACE_CL [*FSS*, 2021]	0.3848 ± 0.0082	0.7538 ± 0.0733	0.5062 ± 0.0235	0.6495 ± 0.0476	0.5209 ± 0.0072	0.0603 ± 0.0034
SCPP-a [*TFS*, 2023]	0.4025 ± 0.0166	0.8065 ± 0.0668	0.5218 ± 0.0260	0.6144 ± 0.0056	0.5535 ± 0.0488	0.2504 ± 0.0123
SCPP-v [*TFS*, 2023]	0.4155 ± 0.0180	0.8111 ± 0.0426	0.5081 ± 0.0197	0.6127 ± 0.0032	0.5572 ± 0.0547	0.2364 ± 0.0134
FMSE-a	$\mathbf{0.4214 \pm 0.0183}$	0.8436 ± 0.0286	$\mathbf{0.6363 \pm 0.0010}$	$\mathbf{0.6910 \pm 0.0443}$	$\mathbf{0.6251 \pm 0.0202}$	$\mathbf{0.2518 \pm 0.0142}$
FMSE-v	$\mathbf{0.4295 \pm 0.0193}$	$\mathbf{0.8562 \pm 0.0162}$	$\mathbf{0.6363 \pm 0.0016}$	0.6791 ± 0.0357	$\mathbf{0.6225 \pm 0.0210}$	0.2480 ± 0.0137

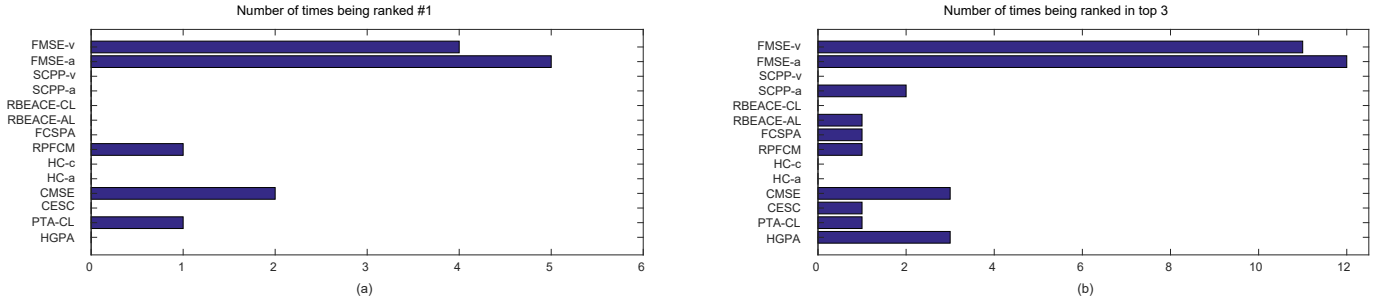


Fig. 3. Number of times that each method is ranked in the (a) first position and (b) top three with respect to Table III.

Selection Database [2] to evaluate the performance of the proposed method, whose details are presented in Table II. These datasets are widely recognized as classical benchmarks for evaluating clustering algorithm performance. Table II provides information on the number of data samples (N), the number of data features (D), and the number of clusters (K).

Base clustering: we conducted 100 and 200 runs of the K-means and FCM clustering algorithms, respectively, to generate base partitions for crisp and fuzzy clustering. To ensure the diversity of the base clustering results, we followed the recommendation in [46]–[48], where k should be greater than the expected number of clusters. We set the number of clusters K for each run of the K-means and FCM algorithms to be randomly selected within the range $[k, \sqrt{n}]$, where k is the true number of clusters, and n is the number of samples. Additionally, to further enhance the diversity of the fuzzy clustering pool, we randomly selected the fuzzy exponent of the FCM algorithm from the interval $[1, 3]$. To provide a fair comparison and reduce the impact of the randomness in the base clustering results, we ran our method multiple times (10 runs) on each dataset, and the average performance was reported.

Evaluation metrics: we employed three widely used external evaluation metrics to evaluate the ensemble clustering results: accuracy (ACC) [49], normalized mutual information (NMI) [6], and adjusted rand index (ARI) [50]. In the experimental analysis, the reference clustering results were the true clustering labels of the datasets. The values of ACC, NMI and ARI range from [0,1], where higher values indicate better clustering performance.

[2]https://jundongl.github.io/scikit-feature/datasets.html

TABLE IV
THE NMI RESULTED FROM DIFFERENT ALGORITHMS, PRESENTED AS "MEAN ± STANDARD DEVIATION"

Method	Seeds	Heart	Ecoli	Climate	Diabetes	WDBC
HGPA [*MLR*, 2002]	0.6871 ± 0.0120	0.0352 ± 0.0075	0.4575 ± 0.0251	0.0039 ± 0.0053	0.0004 ± 0.0007	0.0954 ± 0.0961
PTA-CL [*TKDE*, 2016]	0.4955 ± 0.1138	0.0301 ± 0.0126	0.4830 ± 0.0156	0.0092 ± 0.0125	0.0166 ± 0.0133	0.1967 ± 0.1430
CESC [*CC*, 2021]	0.6195 ± 0.0890	0.0217 ± 0.0152	0.5472 ± 0.0605	0.0094 ± 0.0123	0.0094 ± 0.0095	0.4062 ± 0.0608
CMSE [*TNNLS*, 2023]	0.5723 ± 0.0890	0.0259 ± 0.0164	$\mathbf{0.6541 \pm 0.0470}$	0.0127 ± 0.0095	0.0015 ± 0.0022	0.2071 ± 0.1833
HC-a [*FUZZ-IEEE*, 2014]	0.3662 ± 0.0155	0.0102 ± 0.0072	0.3114 ± 0.0468	0.0094 ± 0.0131	0.0080 ± 0.0072	0.1627 ± 0.0769
HC-c [*FUZZ-IEEE*, 2014]	0.3813 ± 0.0286	0.0158 ± 0.0039	0.3426 ± 0.0403	0.0064 ± 0.0074	0.0072 ± 0.0052	0.1710 ± 0.0382
RPFCM [*FUZZ-IEEE*, 2015]	0.6272 ± 0.0481	0.0199 ± 0.0087	0.3383 ± 0.0388	0.0106 ± 0.0109	0.0004 ± 0.0007	0.4547 ± 0.0148
FCSPA [*AAI*, 2008]	0.6386 ± 0.0219	0.0335 ± 0.0023	0.4391 ± 0.0038	0.0013 ± 0.0000	0.0006 ± 0.0001	0.3828 ± 0.0491
RBEACE_AL [*FSS*, 2021]	0.5964 ± 0.0521	0.0241 ± 0.0169	0.5537 ± 0.0308	0.0067 ± 0.0066	0.0072 ± 0.0078	0.3154 ± 0.1442
RBEACE_CL [*FSS*, 2021]	0.5624 ± 0.0657	0.0254 ± 0.0161	0.4899 ± 0.0134	0.0038 ± 0.0067	0.0020 ± 0.0028	0.3146 ± 0.1301
SCPP-a [*TFS*, 2023]	0.6185 ± 0.0055	0.0201 ± 0.0204	0.5550 ± 0.0282	0.0082 ± 0.0094	0.0066 ± 0.0068	0.3161 ± 0.1495
SCPP-v [*TFS*, 2023]	0.6038 ± 0.0064	0.0231 ± 0.0225	0.5604 ± 0.0234	0.0070 ± 0.0071	0.0067 ± 0.0065	0.3383 ± 0.1549
FMSE-a	$\mathbf{0.7044 \pm 0.0776}$	$\mathbf{0.0359 \pm 0.0097}$	$\mathbf{0.5667 \pm 0.0099}$	$\mathbf{0.0134 \pm 0.0041}$	$\mathbf{0.0204 \pm 0.0134}$	$\mathbf{0.5031 \pm 0.0640}$
FMSE-v	$\mathbf{0.6945 \pm 0.0453}$	$\mathbf{0.0359 \pm 0.0110}$	0.5583 ± 0.0095	$\mathbf{0.0130 \pm 0.0053}$	$\mathbf{0.0202 \pm 0.0138}$	$\mathbf{0.4669 \pm 0.0444}$

Method	CMC	BCW	Waveform	Spambase	SMK	LR
HGPA [*MLR*, 2002]	0.0236 ± 0.0051	0.1753 ± 0.0814	0.2080 ± 0.0676	0.0027 ± 0.0041	0.0432 ± 0.0129	0.3129 ± 0.0076
PTA-CL [*TKDE*, 2016]	0.0227 ± 0.0047	0.1423 ± 0.1406	$\mathbf{0.3948 \pm 0.0181}$	0.0579 ± 0.0318	0.0113 ± 0.0213	$\mathbf{0.4011 \pm 0.0000}$
CESC [*CC*, 2021]	0.0232 ± 0.0076	0.3076 ± 0.1017	0.3602 ± 0.1033	0.0050 ± 0.0112	0.0192 ± 0.0153	0.3109 ± 0.0000
CMSE [*TNNLS*, 2023]	0.0145 ± 0.0087	0.1888 ± 0.1578	$\mathbf{0.4116 \pm 0.0520}$	0.0015 ± 0.0010	0.0209 ± 0.0182	N/A
HC-a [*FUZZ-IEEE*, 2014]	0.0176 ± 0.0070	0.1774 ± 0.0818	0.2587 ± 0.0634	0.0292 ± 0.0197	0.0229 ± 0.0184	0.1149 ± 0.0135
HC-c [*FUZZ-IEEE*, 2014]	0.0262 ± 0.0080	0.2242 ± 0.0593	0.1935 ± 0.0531	0.0198 ± 0.0197	0.0350 ± 0.0252	0.1957 ± 0.0290
RPFCM [*FUZZ-IEEE*, 2015]	0.0249 ± 0.0058	$\mathbf{0.5052 \pm 0.0200}$	0.2861 ± 0.0067	0.0591 ± 0.0026	0.0307 ± 0.0147	0.0491 ± 0.0419
FCSPA [*AAI*, 2008]	0.0215 ± 0.0007	0.3745 ± 0.0202	0.0004 ± 0.0001	0.1114 ± 0.0009	0.0334 ± 0.0036	0.1452 ± 0.0575
RBEACE_AL [*FSS*, 2021]	0.0280 ± 0.0041	0.3451 ± 0.1071	0.2934 ± 0.0168	$\mathbf{0.1118 \pm 0.0132}$	0.0148 ± 0.0038	0.1325 ± 0.0022
RBEACE_CL [*FSS*, 2021]	0.0236 ± 0.0074	0.3205 ± 0.0780	0.2785 ± 0.0230	0.1043 ± 0.0188	0.0134 ± 0.0049	0.1402 ± 0.0102
SCPP-a [*TFS*, 2023]	0.0219 ± 0.0069	0.3262 ± 0.1332	0.2646 ± 0.0761	0.0105 ± 0.0084	0.0213 ± 0.0186	0.3247 ± 0.0066
SCPP-v [*TFS*, 2023]	0.0262 ± 0.0044	0.3322 ± 0.0920	0.2903 ± 0.0628	0.0068 ± 0.0048	0.0210 ± 0.0258	0.3047 ± 0.0190
FMSE-a	$\mathbf{0.0312 \pm 0.0068}$	0.4076 ± 0.0624	0.3062 ± 0.0024	$\mathbf{0.1143 \pm 0.0372}$	$\mathbf{0.0486 \pm 0.0164}$	0.3408 ± 0.0084
FMSE-v	$\mathbf{0.0305 \pm 0.0055}$	$\mathbf{0.4145 \pm 0.0357}$	0.3090 ± 0.0025	0.0805 ± 0.0324	$\mathbf{0.0464 \pm 0.0170}$	$\mathbf{0.3467 \pm 0.0094}$

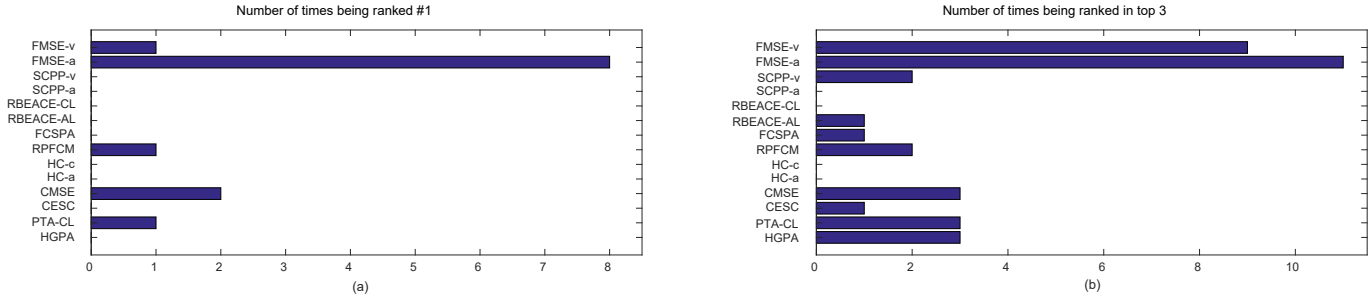


Fig. 4. Number of times that each method is ranked in the (a) first position and (b) top three with respect to Table IV.

B. Clustering Performance Comparison

In this section, we compare the proposed FMSE-a and FMSE-v methods with 12 state-of-the-art clustering ensemble algorithms. The compared methods are four crisp clustering ensemble algorithms (HGPA [6], PTA-CL [51], CESC [14], CMSE [52]), and eight fuzzy clustering ensemble algorithms (FHC-AL [28], FHC-CL [28], RPFCM [32], FCSPA [10], RBEACE-AL [41], RBEACE-CL [41], SCPP-a [15], SCPP-v [15]). Except for the RBEACE algorithm, for which the experimental code was not provided by the original authors, all other compared algorithm codes were supplied by the original authors. The recommended parameter values from the original authors were used for the parameters in the compared algorithms. The fuzzy base clustering adopted by the proposed algorithm is exactly the same as that used by the compared algorithms to ensure the fairness of the experiments.

The average ACC and NMI scores along with their standard deviations for the 14 clustering ensemble algorithms on the 12 datasets are presented in Tables III and IV. The highest index value for each dataset is marked with double underscore bold, the second-highest index value is marked with single underscore bold, and "N/A" denotes that the algorithm encountered a memory shortage error during execution. The ranking information corresponding to Tables III and IV is illustrated in Figs. 3 and 4.

From Table III, our FMSE method achieved the highest ACC scores on the Seeds, Heart, Climate, Diabetes, WDBC, CMC, Spambase, and SMK datasets. For datasets where the highest value was not obtained (e.g., BCW and Waveform), our two methods still secured the second and third positions. From Fig. 3(a), FMSE-a and FMSE-v ranked first in 5 and 4 out of 12 comparisons, respectively, while the third-ranked

method secured the first position only twice. Similarly, From Fig. 3(b), in a total of 12 comparisons, FMSE-a and FMSE-v ranked third in all 12 and 11 cases, respectively, while the third-ranking method CMSE achieved a top three position in only three comparisons, highlighting the significant advantage of FMSE methods in terms of ACC.

With regard to the NMI metric, from Table IV, the FMSE methods achieved the highest NMI scores on the Seeds, Heart, Climate, Diabetes, WDBC, CMC, Spambase, and SMK datasets. From Fig. IV(a), FMSE-a and FMSE-v ranked first in 8 and 1 out of 12 comparisons, respectively. From Fig. IV(b), in a total of 12 comparisons, FMSE-a and FMSE-v achieved top three positions in 11 and 9 cases, respectively, while the third-ranking methods PTA-CL and HGPA secured a top three position in only 3 comparisons. This finding indicates that our FMSE generally outperform the comparison methods in terms of NMI.

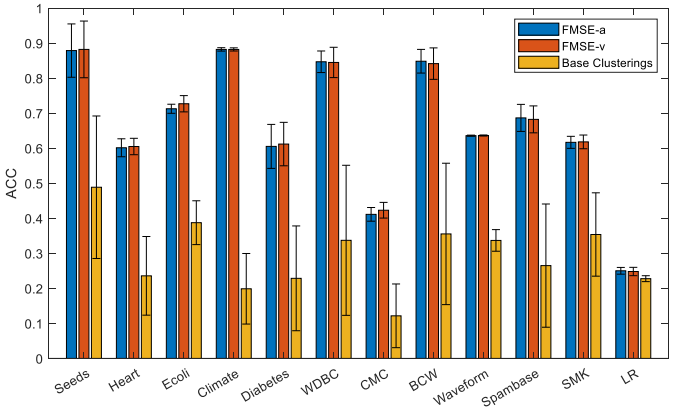


Fig. 5. Average performances in terms of ACC of our methods and the base clustering 100 runs.

Furthermore, our algorithm demonstrates significant advantages over fuzzy clustering ensemble algorithms based on the FCA matrix (e.g., HC-a, HC-c, RBEACE-AL, RBEACE-CL, SCPP-a, SCPP-v). On all 12 datasets, the proposed FMSE-a and FMSE-v methods achieved either the first or second position in all comparisons. This finding strongly supports the notion that our EFCA matrix model can handle sample fuzziness more effectively than the FCA matrix model can, leading to improved clustering results.

Once again, in comparison with crisp clustering ensemble methods, although the CMSE algorithm performed well on the Ecoli and Waveform datasets and the PTA-CL algorithm performed well on the LR dataset, our fuzzy algorithms FMSE-a and FMSE-v achieved better rankings on all other datasets. For instance, on the Climate dataset, FMSE-a and FMSE-v achieved ACC values of 0.8815 and 0.8837, respectively, while the best-performing crisp clustering ensemble algorithm, CESC, obtained a score of only 0.6472. This result clearly demonstrates the superiority of our fuzzy clustering ensemble methods over crisp clustering ensemble methods.

In conclusion, across the 12 datasets from diverse domains and compared with 12 benchmark methods, the proposed FMSE algorithm exhibits the best clustering performance, providing strong evidence of its superiority.

C. Comparison Against Base Clustering

Clustering ensemble aims to integrate multiple base clustering results into a more robust and accurate consensus clustering result. In this section, we compare the consensus clustering results of the proposed FMSE-a and FMSE-v methods with the fuzzy base clustering results. We ran the FMSE methods and the FCM method 100 times on each dataset and calculated the average ACC scores and their error values for the 100 runs. The experimental results are shown in Fig. 5.

Fig. 5 shows that our method significantly improves the ACC scores compared with base clustering on all datasets. For instance, on the Climate dataset, the average ACC score of the base clustering is 0.1987, indicating low clustering quality. However, after base clustering is integrated with our FMSE-a and FMSE-v methods, the average ACC scores improved to 0.8820 and 0.8819, respectively. Similarly, on the WDBC dataset, the ACC scores of the FMSE-a and FMSE-v methods increased from 0.3371 to 0.8467 and 0.8448, respectively, demonstrating a substantial advantage over base fuzzy clustering. The error values calculated on each dataset also further indicate the robustness of our method.

D. Parameter Analysis

Two main hyperparameters are involved in our FMSE algorithm, as shown in Table V. The ranges for the hyperparameters θ and α are $(0, +\infty)$ and $[0, 1)$, respectively. In this section, we further illustrate the impact of θ and α on the FMSE algorithm and provide the parameter ranges for optimal algorithm performance. The experimental results are presented in Fig. 6.

TABLE V
INPUT PARAMETERS OF THE PROPOSED METHOD

Parameter	Description
θ	Parameter that adjusts the fuzziness impact on FRI.
α	Parameter that adjusts the principal components.

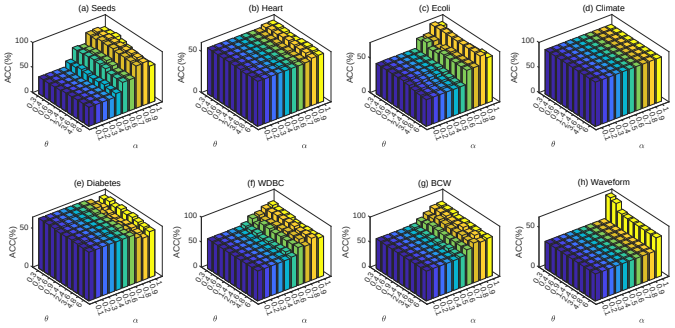


Fig. 6. Clustering performances with respect to ACC with varying θ and α.

Fig. 6 illustrates the ACC performance of the FMSE algorithm with varying values for the hyperparameters θ and α. Each result represents the average value over 30 runs. The x-axis and y-axis denote different values for the parameters θ and α, respectively, while the z-axis represents the final performance evaluation metric, ACC.

First, the graph shows that as θ changes from 0.3 to 6, the corresponding ACC values exhibit no significant variations,

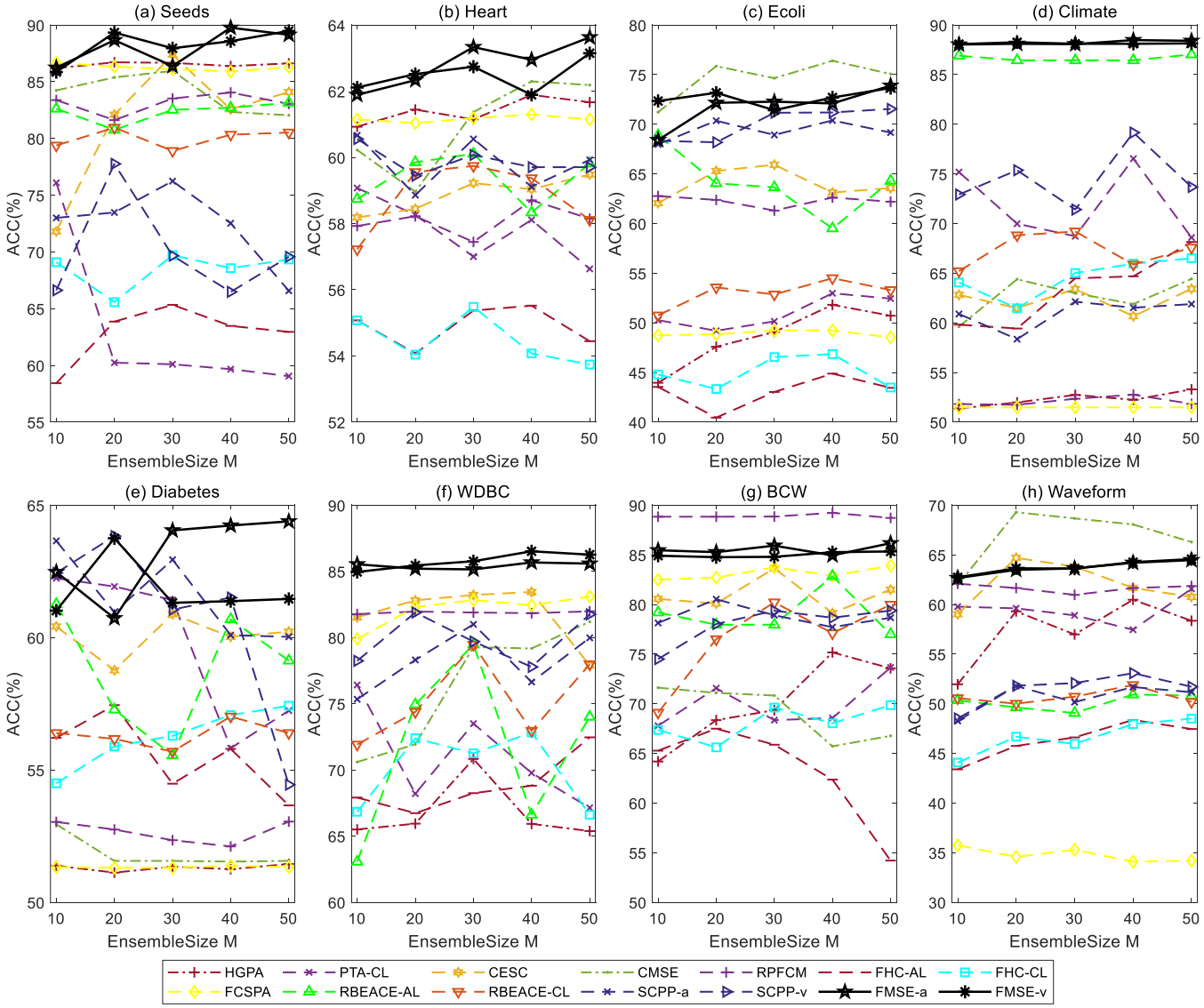


Fig. 7. Clustering performances of different algorithms by varying ensemble size with respect to ACC.

indicating the stability of the parameter θ. On the Seeds dataset, the ACC achieved its highest value when α was in the range $[0.8, 0.9]$. On datasets such as Heart, Ecoli, WDBC, and BCW, the algorithm demonstrated optimal performance when α was set to 0.9. Furthermore, on the Seeds, Ecoli, WDBC, and BCW datasets, the algorithm performance improved when α exceeded 0.7 than when α was in the range $[0, 0.6]$. With all aspects taken into account, the range for α should be set to $[0.7, 1]$, where the FMSE algorithm can achieve optimal performance.

After multiple experiments, this study set the parameter values for the compared algorithms as $\theta = 0.84$ and $\alpha = 0.95$. The FMSE algorithm achieved the best performance under these parameters. This paper presents the experimental results on eight datasets only because of space constraints. Similar parameter analysis results were observed for the remaining datasets.

E. Influence of Ensemble Size

In this section, we evaluate the performance of the FMSE method under different ensemble sizes M. The proposed method is compared with the other 12 ensemble clustering algorithms. For each ensemble size M, all algorithms were run 30 times on each benchmark dataset, and their average performance was recorded. In each run, M base clustering results were randomly selected to participate in the ensemble. The experimental results are depicted in Fig. 7, where the x-axis represents different ensemble sizes M, the y-axis represents the final performance metric ACC, and each curve represents a different method.

As illustrated in Fig. 7, both FMSE algorithms achieved optimal performance on the Seeds, Heart, Climate, Diabetes, and WDBC datasets. The CMSE method demonstrated favorable performance on the Ecoli and Waveforms datasets, yet our methods consistently outperformed CMSE on all other datasets and maintained the second-best position on these two datasets. Moreover, our methods exhibited overall robust

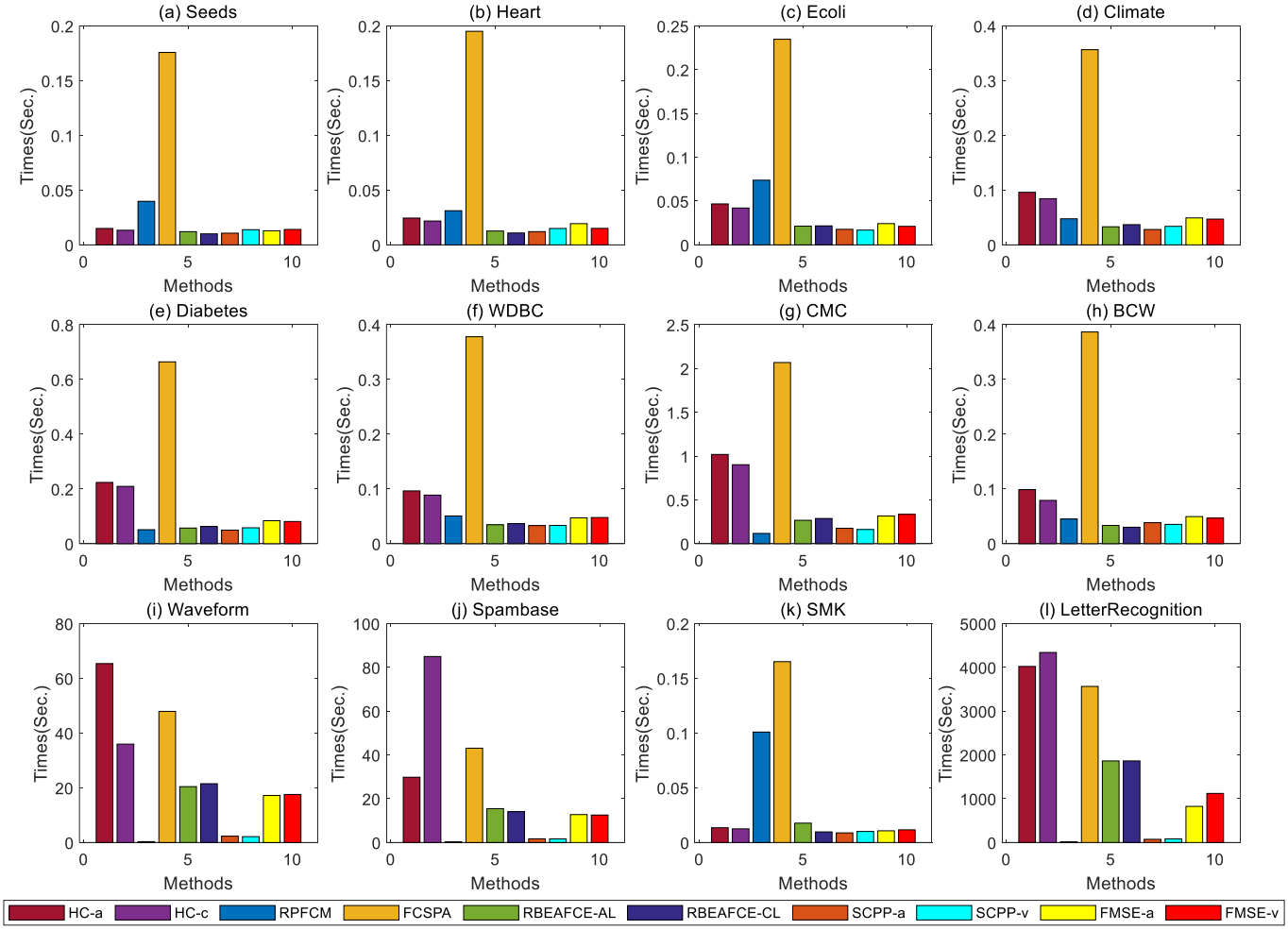


Fig. 8. Running time (seconds) of the different ensemble clustering methods on all data sets.

performance, showing minimal variations under perturbations caused by different ensemble sizes (as observed on WDBC and Waveforms, where our algorithm's performance steadily improved with the increase in M). This finding indicates that our method possesses good robustness. Remarkably, our algorithm outperformed most of the methods, even with an ensemble size as small as 10, while others used 50 base clustering as input.

F. Execution Time

In this section, we evaluated the average execution time of the proposed FMSE and all fuzzy clustering ensemble algorithms over 10 runs on all benchmark datasets, and the experimental results are shown in Fig.8. Overall, the time cost of FMSE is at a moderate level, which is equivalent to the efficiency of most clustering ensemble algorithms based on fuzzy matrix. Even on the largest dataset LR, FMSE is consistently faster than most fuzzy matrix-based methods (e.g. FCSPA, HC-a, HC-c). Although our running speed is slower than SCPP, considering that our method significantly outperforms SCPP in clustering performance, a little sacrifice in running time is acceptable.

G. Ablation Study

In this section, we investigate the necessity of the two strategies involved in the proposed FMSE algorithm when enhancing the FCA matrix. First, we abandon the SVD strategy and name this degenerate version "w/o SVD". Second, we remove the influence of FRI on the FCA matrix and name it "w/o FRI".

Table VI demonstrates the clustering performance of FMSE after removing the above-mentioned two strategies. It can be observed that FMSE with both strategies retained consistently achieves the best experimental results across most benchmark datasets. Specifically, for the degraded version without SVD (w/o SVD(a)), it lags behind in 32 out of 36 comparisons with FMSE-a, indicating the necessity of the SVD strategy for the proposed algorithm. Similarly, for the degraded version without FRI (w/o FRI(a)), it only achieves victory in 2 out of 36 comparisons, demonstrating the essential role of the FRI strategy. Similar observations can be made in the ablation study of FMSE-v as well. In summary, both strategies involved in the proposed algorithm are indispensable for improving the performance of the FMSE algorithm.

V. Conclusion

The current fuzzy clustering ensemble methods are mostly based on FCA matrices. However, the low-value density and

TABLE VI
ABLATION STUDY IN ACC, NMI AND ARI FOR THE PROPOSED FMSE-A

Metric	ACC			NMI			ARI		
Algorithm	w/o SVD	w/o FRI	FMSE-a	w/o SVD	w/o FRI	FMSE-a	w/o SVD	w/o FRI	FMSE-a
Seeds	0.8349	0.8159	**0.8875**	0.6824	0.6676	**0.7044**	0.6549	0.6059	**0.7202**
Heart	0.5686	0.6095	**0.6220**	0.0276	0.0302	**0.0359**	0.0159	0.0437	**0.0457**
Ecoli	0.5735	0.7101	**0.7116**	0.5060	**0.5724**	0.5667	0.3927	0.5670	**0.6138**
Climate	0.6274	0.8814	**0.8815**	**0.0160**	0.0127	0.0134	-0.0077	0.0014	**0.0303**
Diabetes	**0.6285**	0.5709	0.6206	0.0164	0.0107	**0.0204**	0.0448	0.0276	**0.0536**
WDBC	0.7762	0.8216	**0.8487**	0.3541	0.4437	**0.5031**	0.3140	0.4283	**0.4938**
CMC	0.4047	0.4151	**0.4214**	0.0251	0.0309	**0.0312**	0.0191	0.0228	**0.0263**
BCW	0.8221	0.8229	**0.8436**	0.3583	0.4017	**0.4076**	0.3954	0.4316	**0.5103**
Waveform	0.5055	0.6354	**0.6363**	**0.3776**	0.3072	0.3062	0.2912	0.3642	**0.3653**
Spambase	0.6432	0.6804	**0.6910**	0.0229	0.1018	**0.1143**	0.0199	0.1064	**0.1361**
SMK	0.6113	0.6122	**0.6251**	0.0408	0.0411	**0.0486**	0.0356	0.0358	**0.0532**
LR	0.2491	0.2494	**0.2518**	0.3328	**0.3460**	0.3408	**0.1261**	0.1199	0.1128

uniform dispersion of FCA matrix significantly affect clustering performance. To address this issue, we introduce SVD to alleviate the low-value density of FCA and simultaneously mitigate its uniform dispersion through fuzzy representative-weighted structures of the samples. On this basis, we employ a consensus method using self co-association prototype diffusion to obtain the final ensemble clustering result. Extensive experimental comparisons validate the superiority of the proposed FMSE algorithm.

This paper utilizes fuzzy matrix to capture the co-occurrence pattern of pairs of samples while considering the fuzziness of the samples. However, the complexity of the fuzzy matrix leads to high computational costs for the clustering ensemble algorithm based on it, making scalability to large datasets difficult. In the future, we will focus on this scalability issue.

ACKNOWLEDGMENTS

This work was supported by Natural Science Foundation of China under Grant 62176001, Grant 62076002, and, in part by the Natural Science Foundation of Anhui Province under Grant 2008085MF194 and 2108085MF212, and in part by the Key Project of Natural Science Foundation of Anhui Provincial Department of Education under Grant KJ2020A0041 and KJ2021A0043. This work was also supported by Natural Science Project of Anhui Provincial Education Department grants 2023AH030004.

REFERENCES

[1] H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE Transactions on knowledge and data engineering*, vol. 17, no. 4, pp. 491–502, 2005.

[2] N. Iam-On, T. Boongeon, S. Garrett, and C. Price, "A link-based cluster ensemble approach for categorical data clustering," *IEEE Transactions on knowledge and data engineering*, vol. 24, no. 3, pp. 413–425, 2010.

[3] J. P. Sarkar, I. Saha, and U. Maulik, "Rough possibilistic type-2 fuzzy c-means clustering for mr brain image segmentation," *Applied Soft Computing*, vol. 46, pp. 527–536, 2016.

[4] B. Jiang, X. Wu, X. Zhou, Y. Liu, A. G. Cohn, W. Sheng, and H. Chen, "Semi-supervised multiview feature selection with adaptive graph learning," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[5] V. Y. Kiselev, T. S. Andrews, and M. Hemberg, "Challenges in unsupervised clustering of single-cell rna-seq data," *Nature Reviews Genetics*, vol. 20, no. 5, pp. 273–282, 2019.

[6] A. Strehl and J. Ghosh, "Cluster ensembles—a knowledge reuse framework for combining multiple partitions," *Journal of machine learning research*, vol. 3, no. Dec, pp. 583–617, 2002.

[7] X. Zhang, L. Jiao, F. Liu, L. Bo, and M. Gong, "Spectral clustering ensemble applied to sar image segmentation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 7, pp. 2126–2136, 2008.

[8] Z. Yu, L. Li, J. Liu, J. Zhang, and G. Han, "Adaptive noise immune cluster ensemble using affinity propagation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 12, pp. 3176–3189, 2015.

[9] E. Ramasso, V. Placet, and M. L. Boubakar, "Unsupervised consensus clustering of acoustic emission time-series for robust damage sequence estimation in composites," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 12, pp. 3297–3307, 2015.

[10] K. Punera and J. Ghosh, "Consensus-based ensembles of soft clusterings," *Applied Artificial Intelligence*, vol. 22, no. 7-8, pp. 780–810, 2008.

[11] A. L. Fred and A. K. Jain, "Combining multiple clusterings using evidence accumulation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 6, pp. 835–850, 2005.

[12] N. Iam-On, T. Boongoen, S. Garrett, and C. Price, "A link-based approach to the cluster ensemble problem," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 12, pp. 2396–2409, 2011.

[13] F. Li, Y. Qian, J. Wang, C. Dang, and L. Jing, "Clustering ensemble based on sample's stability," *Artificial Intelligence*, vol. 273, pp. 37–55, 2019.

[14] X. Ji, S. Liu, P. Zhao, X. Li, and Q. Liu, "Clustering ensemble based on sample's certainty," *Cognitive Computation*, vol. 13, no. 4, pp. 1034–1046, 2021.

[15] F. Li, J. Wang, Y. Qian, G. Liu, and K. Wang, "Fuzzy ensemble clustering based on self co-association and prototype propagation," *IEEE Transactions on Fuzzy Systems*, 2023.

[16] X. Z. Fern and C. E. Brodley, "Solving cluster ensemble problems by bipartite graph partitioning," in *Proceedings of the twenty-first international conference on Machine learning*, 2004, p. 36.

[17] D. Huang, J.-H. Lai, and C.-D. Wang, "Combining multiple clusterings via crowd agreement estimation and multi-granularity link analysis," *Neurocomputing*, vol. 170, pp. 240–250, 2015.

[18] D. Huang, C.-D. Wang, J.-S. Wu, J.-H. Lai, and C.-K. Kwoh, "Ultra-scalable spectral clustering and ensemble clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 6, pp. 1212–1226, 2019.

[19] P. Zhou, L. Du, and X. Li, "Self-paced consensus clustering with bipartite graph," in *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, 2021, pp. 2133–2139.

[20] D. Huang, J. Lai, and C.-D. Wang, "Ensemble clustering using factor graph," *Pattern Recognition*, vol. 50, pp. 131–142, 2016.

[21] P. Zhou, L. Du, X. Liu, Y.-D. Shen, M. Fan, and X. Li, "Self-paced clustering ensemble," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 4, pp. 1497–1511, 2020.

[22] P. Zhou, L. Du, and X. Li, "Adaptive consensus clustering for multiple k-means via base results refining," *IEEE Transactions on Knowledge and Data Engineering*, 2023.

[23] P. Zhou, B. Hu, D. Yan, and L. Du, "Clustering ensemble via diffusion on adaptive multiplex," *IEEE Transactions on Knowledge and Data Engineering*, 2023.

[24] P. Zhou, L. Du, X. Liu, Z. Ling, X. Ji, X. Li, and Y.-D. Shen, "Partial clustering ensemble," *IEEE Transactions on Knowledge and Data Engineering*, 2023.

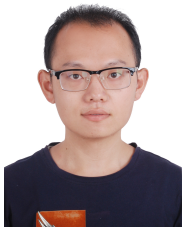
[25] A. Bagherinia, B. Minaei-Bidgoli, M. Hossinzadeh, and H. Parvin, "Elite fuzzy clustering ensemble based on clustering diversity and quality measures," *Applied Intelligence*, vol. 49, pp. 1724–1747, 2019.
[26] C. Domeniconi and M. Al-Razgan, "Weighted cluster ensembles: Methods and analysis," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 2, no. 4, pp. 1–40, 2009.
[27] S.-o. Abbasi, S. Nejatian, H. Parvin, V. Rezaie, and K. Bagherifard, "Clustering ensemble selection considering quality and diversity," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 1311–1340, 2019.
[28] P. Su, C. Shang, and Q. Shen, "Link-based pairwise similarity matrix approach for fuzzy c-means clustering ensemble," in *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2014, pp. 1538–1544.
[29] G. H. Golub and C. Reinsch, "Singular value decomposition and least squares solutions," in *Handbook for Automatic Computation: Volume II: Linear Algebra*. Springer, 1971, pp. 134–151.
[30] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.
[31] L. I. Kuncheva and D. P. Vetrov, "Evaluation of stability of k-means cluster ensembles with respect to random initialization," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 11, pp. 1798–1808, 2006.
[32] M. Popescu, J. Keller, J. Bezdek, and A. Zare, "Random projections fuzzy c-means (rpfcm) for big data clustering," in *2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2015, pp. 1–6.
[33] R. Avogadri and G. Valentini, "Fuzzy ensemble clustering based on random projections for dna microarray data analysis," *Artificial Intelligence in Medicine*, vol. 45, no. 2-3, pp. 173–183, 2009.
[34] M. Ye, W. Liu, J. Wei, X. Hu *et al.*, "Fuzzy-means and cluster ensemble with random projection for big data clustering," *Mathematical Problems in Engineering*, vol. 2016, 2016.
[35] I. S. Dhillon, S. Mallela, and R. Kumar, "A divisive information theoretic feature clustering algorithm for text classification," *The Journal of machine learning research*, vol. 3, pp. 1265–1287, 2003.
[36] S. Kullback and R. A. Leibler, "On information and sufficiency," *The annals of mathematical statistics*, vol. 22, no. 1, pp. 79–86, 1951.
[37] P. Zhou, B. Sun, X. Liu, L. Du, and X. Li, "Active clustering ensemble with self-paced learning," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.
[38] M. Mojarad, S. Nejatian, H. Parvin, and M. Mohammadpoor, "A fuzzy clustering ensemble based on cluster clustering and iterative fusion of base clusters," *Applied Intelligence*, vol. 49, pp. 2567–2581, 2019.
[39] E. Bedalli, E. Mançellari, and O. Asilkan, "A heterogeneous cluster ensemble model for improving the stability of fuzzy cluster analysis," *Procedia Computer Science*, vol. 102, pp. 129–136, 2016.
[40] V. Berikov, "A probabilistic model of fuzzy clustering ensemble," *Pattern Recognition and Image Analysis*, vol. 28, pp. 1–10, 2018.
[41] A. Bagherinia, B. Minaei-Bidgoli, M. Hosseinzadeh, and H. Parvin, "Reliability-based fuzzy clustering ensemble," *Fuzzy Sets and Systems*, vol. 413, pp. 1–28, 2021.
[42] Y. Zhang, F.-L. Chung, and S. Wang, "A multiview and multiexemplar fuzzy clustering approach: theoretical analysis and experimental studies," *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 8, pp. 1543–1557, 2018.
[43] Z. Kang, Z. Lin, X. Zhu, and W. Xu, "Structured graph learning for scalable subspace clustering: From single view to multiview," *IEEE Transactions on Cybernetics*, vol. 52, no. 9, pp. 8976–8986, 2021.
[44] Z. Lin, Z. Kang, L. Zhang, and L. Tian, "Multi-view attributed graph clustering," *IEEE Transactions on knowledge and data engineering*, vol. 35, no. 2, pp. 1872–1880, 2021.
[45] S. J. Choudhury and N. R. Pal, "Fuzzy clustering of single-view incomplete data using a multiview framework," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 12, pp. 5312–5323, 2022.
[46] H. G. Ayad and M. S. Kamel, "Cumulative voting consensus method for partitions with variable number of clusters," *IEEE transactions on pattern analysis and machine intelligence*, vol. 30, no. 1, pp. 160–173, 2007.
[47] L. I. Kuncheva and S. T. Hadjitodorov, "Using diversity in cluster ensembles," in *2004 IEEE international conference on systems, man and cybernetics (IEEE Cat. No. 04CH37583)*, vol. 2. IEEE, 2004, pp. 1214–1219.
[48] Ayad, Hanan G and Kamel, Mohamed S, "On voting-based consensus of cluster ensembles," *Pattern Recognition*, vol. 43, no. 5, pp. 1943–1953, 2010.
[49] Y. Yang, "An evaluation of statistical approaches to text categorization," *Information retrieval*, vol. 1, no. 1-2, pp. 69–90, 1999.
[50] L. Hubert and P. Arabie, "Comparing partitions," *Journal of classification*, vol. 2, pp. 193–218, 1985.
[51] D. Huang, J.-H. Lai, and C.-D. Wang, "Robust ensemble clustering using probability trajectories," *IEEE transactions on knowledge and data engineering*, vol. 28, no. 5, pp. 1312–1326, 2015.
[52] Y. Jia, S. Tao, R. Wang, and Y. Wang, "Ensemble clustering via co-association matrix self-enhancement," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

Xia Ji received her BS and PhD degrees in Computer Science from Anhui University in 2005 and 2010. She is currently an associate professor at the School of Computer Science and Technology, Anhui University. Her research interests include artificial intelligence, machine learning, data mining, and intelligent information processing.

Jiawei Sun is currently studying for a master's degree in School of Computer Science and Technology, Anhui University. His research interests include multi-view clustering and clustering ensemble.

Jianhua Peng is currently studying for a master's degree in School of Computer Science and Technology, Anhui University. His research interests include clustering ensemble.

Yue Pang is currently studying for a master's degree in School of Computer and Information Engineering, Henan University. Her research interests include clustering ensemble.



Peng Zhou received the B.E. degree in computer science and technology from University of Science and Technology of China in 2011, and Ph.D degree in Computer Science from Institute of Software, Chinese Academy of Sciences in 2017. He is currently an Associate Professor with the School of Computer Science and Technology, Anhui University. He has published more than 40 papers in highly regarded conferences and journals, including IEEE TKDE, IEEE TNNLS, IEEE TCYB, ACM TKDD, IJCAI, AAAI, MM, SDM, ICDM, etc. His research interests include machine learning, data mining and artificial intelligence.