IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 11 January 2024, accepted 24 January 2024, date of publication 29 January 2024, date of current version 9 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3359299

RESEARCH ARTICLE

Fuzzy Divergence Weighted Ensemble Clustering With Spectral Learning Based on Random Projections for Big Data

INES LAHMAR [1], AIDA ZAIER [2], MOHAMED YAHIA [3], TARIG ALI [4], (Senior Member, IEEE), AND RIDHA BOAULLEGUE [2]

[1]MACS Laboratory, University of Gabes, Gabes 6029, Tunisia
[2]Innov'Com Laboratory, University of Carthage Tunis, Tunis 1002, Tunisia
[3]SYSCOM Laboratory, ENIT, Université Tunis El Manar, Tunis 1002, Tunisia
[4]GIS and Mapping Laboratory, American University of Sharjah, Sharjah, United Arab Emirates

Corresponding author: Ines Lahmar (ines12lahmar@gmail.com)

This work was supported by the Deanship of Scientific Research (DGRST), Tunisia, under the ARUB Program.

ABSTRACT In many real-world applications, data are described by high-dimensional feature spaces, posing new challenges for current ensemble clustering methods. The goal is to combine sets of base clusters to enhance clustering accuracy, but this makes them susceptible to low quality. However, the reliability of present ensemble clustering in high-dimensional data still needs improvement. In this context, we propose a new fuzzy divergence-weighted ensemble clustering based on random projection and spectral learning. Firstly, random projection (RP) is used to create various dimensional data and find membership matrices via fuzzy c-means (FCM). Secondly, fuzzy partitions of random projections are ranked using entropy-based local weighting along with Kullback-Leibler (KL) divergence to detect any uncertainty. Then it used to evaluate the weight of each cluster. Finally, we create regularized graphs from these membership matrices and use spectral matrices to estimate the affinity matrices of these graphs using fuzzy KL divergence anchor graphs. Subsequently, obtaining the final clustering results is considered as an optimization problem, and the ensemble clustering results are obtained. The experimental results on high-dimensional data demonstrate the efficiency of our method compared to state-of-the-art methods.

INDEX TERMS Fuzzy ensemble clustering, high-dimensional data, random projection, Kullback-Leibler divergence entropy, spectral learning.

I. INTRODUCTION

Clustering, an unsupervised learning algorithm, is employed to identify inherent structures within unlabeled data, manifesting as clusters. The surge in data streaming from diverse sources has led to an influx of high-dimensional data, presenting a challenge. This challenge can be addressed by reducing the data volume to a manageable level while preserving essential information [1].

Recently, clustering ensemble algorithms have become an important task for clustering analysis. They are based on the use of multiple base clusters to generate stable consensus clustering. Ensemble clustering can typically be divided into two phases: cluster generation and consensus clustering. An intuitive method to create low-quality base clustering has been developed. Some methods have been developed to assess each ensemble member and assign it a weight to enhance consensus clustering [2]. Li et al. designed multi-view clustering via a learned bipartite graph [3]. Huang et al. proposed the local diversity strategy of clusters inside a base clustering structure [4]. Rathore et al. proposed fuzzy ensemble clustering using cumulative aggregation on random projections [5].

In fuzzy clustering ensembles, researchers have proposed consensus methods to address challenges related to the alignment of membership matrices from base clustering.

The associate editor coordinating the review of this manuscript and approving it for publication was Alberto Cano.

To achieve consensus, the former either performs column exchanges to complete the alignment or converts the membership matrices of the data graphs into affinity matrices (for example, the co-similarity matrix) [6]. Recently, numerous graph-based clustering ensembles have been proposed [7]. However, these methods construct the co-similarity matrix by multiplying the membership matrix and its transposed form [8]. Consequently, information is lost, leading to the creation of an erroneous affinity graph. Moreover, merging large-scale graphs requires significant storage space and processing power. Currently, the authors suggest employing spectral matrices to scale down graphs [9]. Additionally, column permutation has been demonstrated to be ineffective for random projection (RP) and fuzzy c-means (FCM) [10].

For high-dimensional data, clustering ensembles face two challenges [11]. Firstly, the curse of dimensionality is introduced by the insufficient number of samples, making it challenging to find statistically meaningful structures in the data. Secondly, irrelevant features can degrade the performance of ensemble clustering. Dimensionality reduction is a method that could be employed to enhance the utility of clustering for high-dimensional data. Numerous researchers have proposed ensemble clustering approaches for high-dimensional data, leading to the development of feature-based models such as random subspace [12], feature sampling [13], and random projection (RP) [5], [11].

In this paper, we propose a fuzzy ensemble clustering based on random projection with fuzzy c-means (RPFCM) to aggregate fuzzy membership matrices for high-dimensional data. Furthermore, to achieve cluster-wise weighted diversity in multiple base clustering, we develop an entropy-based cluster validity measure by combining entropy with Kullback-Leibler (KL) divergence to locally weight the clusters. These measures are used to determine the reliability of consensus partitions. After obtaining locally weighted divergence, we convert the membership matrices of each base clustering into regularized graphs and use spectral matrices to approximate the affinity matrices using KL divergence anchor graphs. The process of integrating the matrices into the final result is regarded as an optimization problem. Multiple experiments are conducted on high-dimensional data, and the results indicate the efficiency of the proposed ensemble clustering algorithm compared to state-of-the-art algorithms.

Here is an outline of the remainder of this paper. Section II covers related works. Section III introduces the proposed method. Section IV presents the experimental results and comparisons. Finally, Section V concludes the paper.

II. RELATED WORK

Ensemble clustering for high-dimensional data poses unique challenges, including the curse of dimensionality and the influence of irrelevant features. To address these challenges and enhance the utility of clustering in high-dimensional spaces, researchers have explored dimensionality reduction techniques. Feature-based models, such as RP [5], [11], random subspace [12], and feature sampling [13], have been proposed to improve the efficiency and effectiveness of clustering algorithms when applied to datasets with a high number of dimensions.

As a consequence, various FCM ensemble methods [2], employing shallow randomization-based techniques like RP, have been investigated. Among them, RP-based FCM clustering (RPFCM) [6] has garnered interest. RP is a relatively simple model that only requires creating a random matrix according to a preset distribution function, contrasting with traditional feature extraction methods that necessitate optimization algorithms. RP distance preservation property ensures minimal differences between pairwise distances in the original and low-dimensional spaces, making it practical for coupling with distance-based FCM.

The work in FCM ensembles can be broadly classified into two groups based on their objectives. The first is the data-based approach, and the second is the clustering-based approach.

In fuzzy clustering ensemble methods, a significant challenge arises from aligning membership matrices from base clustering [5]. Researchers have introduced consensus and cascade methods to mitigate this challenge. The former involves column exchanges for alignment or converts membership matrices into affinity matrices of data graphs [10], such as graph-based clustering ensemble frameworks [7] and the co-similarity matrix [8].

III. PROPOSED METHOD

This part will introduce the prediction method of divergence-based locally weighted RPFCM ensemble clustering, aiming to transform multi-base clustering into a more robust consensus clustering for the final result. In this section, we elaborate on each step of the method in detail. First, the RP method is introduced in Section A. Next, the entropy-based locally weighted method is presented in Section B. Then, the KL divergence-based locally weighted method is described in Section C. The fuzzy divergence anchor graph is outlined in Section D. The base clustering generation based on divergence entropy with RPFCM is detailed in Section E. Finally, fuzzy consensus spectral learning and an iterative optimization method are explained in Section F. The flowchart of the proposed method is depicted in FIGURE 1.

A. RANDOM PROJECTION (RP)

RP is a widely used method for reducing the dimensionality of high-dimensional data. RP involves a linear transformation from $\mathbb{R}^p$ to a lower-dimensional space $\mathbb{R}^q$, which is described by a matrix R. Given a set of data points $X = \{x_1, x_2, \ldots, x_n\} \subset \mathbb{R}^p$, representing a dataset in the "upspace," RP allows us to convert X into a reduced-dimensional dataset $Z = \{z_1, z_2, \ldots, z_n\} \subset \mathbb{R}^q$, referred to as the "downspace," through the linear

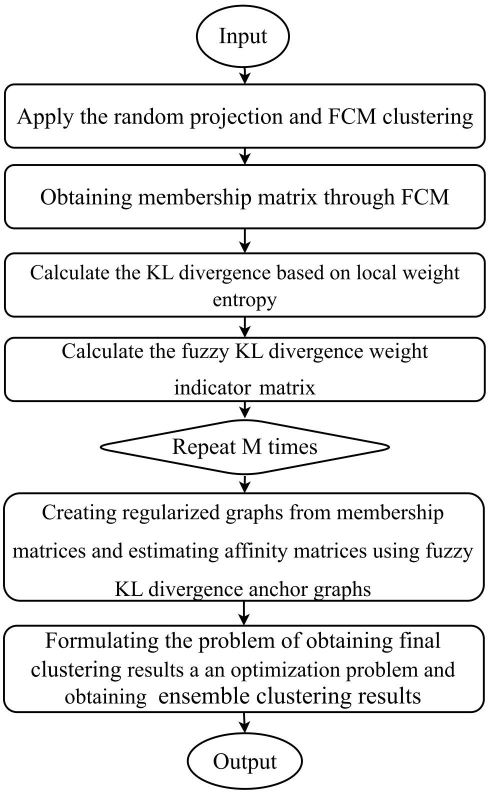


FIGURE 1. Flow diagram of the proposed method.

transformation of X using R [5]. Most RP methods are based on the Johnson-Lindenstrauss (JL) lemma [14].

Lemma: Given a matrix $X \in \mathbb{R}^{n \times p}$ that represents a dataset of n data points in p dimensions, denoted as the "upspace," and let $\varepsilon, \beta > 0$, for any integer q satisfying:

$$q \geq q_0 = \frac{4 + 2\beta}{\varepsilon^2/2 - \varepsilon^3/3} \log n. \tag{1}$$

q_0 is the lower bound on the dimensionality of the reduced space. The factor ε controls the accuracy of distance preservation, while β controls the probability that distance preservation to within $1 \pm \varepsilon$ is achieved. Then, for any set Z of n data points $\in \mathbb{R}^q$, denoted as the "downspace," there exists a linear map $f : \mathbb{R}^p \to \mathbb{R}^q$. For any $u, k \in Z$, with probability at least $1 - n^{-\beta}$, the dimensionality of the low-dimensional space should satisfy:

$$(1-\varepsilon)||x_i - x_j||^2 \leq ||f(x_i) - f(x_j)||^2 \leq (1+\varepsilon)||x_i - x_j||^2. \tag{2}$$

Lemma implies that if the reduced dimension q is equal or bigger than the JL lower bound q_0, then pairwise Euclidean distance squares are preserved within a multiplicative factor of 1ϵ.

Random projection theory in RPFCM involves several important steps. First, it guides the construction of the RP matrix R by assigning values $(+1, 0$ or $-1))$ to its elements. This process lets that R retains the important property of maintaining the connection from high to low. This matrix R is then used to perform a linear transformation and projects the dataset into the reduced space $\mathbb{R}^q$ as shown below:

$$X_l = \frac{1}{\sqrt{q}} R^T X. \tag{3}$$

R is selected to be a Gaussian random matrix, but the authors proposed a database-friendly RP to replace it with a simple and sparse matrix R [15], which can be calculated.

$$R = \sqrt{3} \times \begin{cases} +1 & \text{with probability} = \dfrac{1}{6} \\ 0 & \text{with probability} = \dfrac{2}{3} \\ -1 & \text{with probability} = \dfrac{1}{6} \end{cases} \tag{4}$$

Note that Eq. (2) provides that the low-dimensional data X_l produced by Eq. (3) respects the pairwise distances in the high-dimensional space to a certain extent. The error between the results of performing FCM on X_l and on X is constrained within a certain range. This ensures that RPFCM is theoretically feasible.

B. ENTROPY-BASED LOCALLY WEIGHTED

Suppose there is a dataset $X = \{x_1, \ldots, x_n\}$ with n objects. The dataset X is replicated M times to obtain M partitions, where each partition contains a certain number of clusters. The set Π of M-based clusters is as follows:

$$\Pi = \left\{\Pi^1, \Pi^2, \ldots, \Pi^M\right\}. \tag{5}$$

where

$$\Pi^m = \left\{c_1^m, c_2^m, \ldots, c_{n^m}^m\right\}. \tag{6}$$

where Π^m is the m-th base cluster in Π. Each base cluster consists of a number of clusters.

c_j^m is the j-th cluster, and n^m indicates the number of clusters in the m-th base cluster Π^m. All clusters in the base cluster set Π are denoted as follows:

$$C = \left\{c_1, c_2, \ldots, c_{nc}\right\}. \tag{7}$$

where c_i is the i-th base clustering, and nc is the total number of clusters in Π.

Each cluster is a set of data objects. To estimate the reliability of each cluster, the entropy-based locally weighted measure calculates the uncertainty of the cluster by considering the cluster labels in the entire set. The entropy of a cluster, $c_i \in C$, for the base cluster set Π, can be calculated [4]:

$$E^{\Pi}(c_i) = \sum_{m=1}^{M} - \sum_{j=1}^{n^m} p_{entropy}(c_i, c_j^m) \log_2 p_{entropy}(c_i, c_j^m). \tag{8}$$

With

$$p_{entropy}(c_i, c_j^m) = \frac{|c_i \cap c_j^m|}{|c_i|}. \tag{9}$$

where $\cap$ computes the intersection of two sets (or clusters). M denotes the number of base clusters. $|c_i|$ denotes the number of elements in the cluster Ci.

As it stands, $p_{entropy}(c_i, c_j^m)$ is in the interval [0, 1] for any i, j, and m, resulting in $E^{\Pi}(c_i) \in [0, \infty)$. The uncertainty of c_i with respect to Π is zero when all of data objects in c_i belong to the same fuzzy cluster in Π.

After calculating the entropy of each cluster, we consider the uncertainty of the cluster relative to the set through the notion of ECI and add weights to the data items within each cluster.

Given a cluster set of M base clusters, the ECI of each cluster is computed as follows:

$$ECI(c_i) = \exp^{-\frac{E^{\Pi}(c_i)}{\theta \cdot M}}. \tag{10}$$

where $\theta > 0$ is the parameter used to adjust the degree of fuzzy cluster unreliability with respect to the index. After obtaining the entropy of each cluster, it is applied to the selected similarity matrix to generate the final data matrix.

C. KL DIVERGENCE-BASED LOCALLY WEIGHTED ENTROPY METHOD

To uncover any uncertainty or randomness within a cluster, we combine fuzzy entropy with the KL divergence method. The accuracy of clustering results hinges on understanding how to evaluate differences between distributions. When two distributions are equivalent, the KL divergence is 0. We propose utilizing KL divergence as a measure of local weights. It's important to note that the KL divergence results in this study are not always guaranteed to be greater than 0, as $p_{KL}(c_i, c_j^m)$ and $p_{KL}(c_j^m, c_i)$ lack straightforward probabilistic interpretations. KL divergence weights are calculated as follows:

$$Kl^{\Pi}(c_i) = \sum_{m=1}^{M}\sum_{j=1}^{N} p_{KL}(c_i, c_j^m)\log_2 \frac{p_{KL}(c_i, c_j^m)}{p_{KL}(c_j^m, c_i)}. \tag{11}$$

Here, $p_{KL}(c_i, c_j^m) = \frac{|c_i \cap c_j^m|}{|c_i|}$, where $|c_i \cap c_j^m|$ represents the number of elements in the intersection between cluster c_i and data point c_j^m, and N is the number of clusters in the partition Π^m.

After obtaining the KL divergence for each cluster in the set, we assess the uncertainty of the cluster concerning the entire set through the DECI and assign weights to the data objects within each cluster.

$$DECI(c_i) = \exp^{-\frac{Kl^{\Pi}(c_i)}{\theta \cdot M}}. \tag{12}$$

Here, $Kl^{\Pi}(c_i)$ represents the KL divergence weight for cluster c_i, $\theta > 0$ is the parameter adjusting the level of fuzziness in cluster unreliability, and M is the total number of clusters in the dataset.

Currently, for each $c_i \in C$, $DECI(c_i)$ falls within the range (0, 1]. When the entropy of a cluster reaches its minimum value ($Kl^{\Pi}(c_i) = 0$), its DECI reaches its maximum value ($DECI(c_i) = 1$). The DECI of a cluster tends to approach zero as the entropy of the cluster approaches infinity.

D. FUZZY DIVERGENCE ANCHOR GRAPH

A graph $G = \{V, E, W\}$ is constructed by the set of vertices V, the set of edges E, and the weight matrix W. $W = [w_{ij}]_{N\times N}$, also known as the affinity matrix, specifies the similarity of two data samples. In general, the KNN and anchor graph can be used to construct the matrix W [16].

In large-scale datasets, it is difficult to build the affinity matrix W directly. To resolve this problem, we apply the anchor-based graph, which has a powerful construction method. In general, the anchor graph seeks M anchor points from the original data samples.

Let M anchors $P = [p_1, p_2, \ldots p_M]$, we can represent pairwise relationships between the i-th data sample and j-th anchor samples using a bipartite graph with a weight matrix, which can be defined as follows:

$$Q_{ij} = \begin{cases} \dfrac{DECI(c_i)}{\sum_{j'} DECI(c_i)} & \text{if } j \in < i > \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

where $< i >$ represents the r nearest anchor point to c_i. DECI represents the divergence-based locally weighted. From the bipartite graph, the full affinity matrix W can be defined as follows:

$$W = Q[Diag(1_N^T Q)]^{-1}Q^T = QA^{-1}Q^T. \tag{14}$$

In the fuzzy anchor graph, prototype C can be thought of as the anchors, and prototype U can be thought of as the weight matrix of the anchor graph. Then, for M membership matrices $\{U_{(i)}\}_{i=1}^{M}$, we can convert them into corresponding affinity matrices $\{A_{(i)}\}_{i=1}^{M}$ according to the following equation:

$$A_{(i)} = U_{(i)}A^{-1}U_{(i)}^T; \quad i = 1, 2, \ldots, M. \tag{15}$$

According to the affinity matrix $A_{(i)}$, its Laplacian matrix L for performing dimensionality reduction on the data can be defined as follows:

$$L(i) = \overline{A_{(i)}} - D(i). \tag{16}$$

where D is a diagonal matrix representing the degree matrix $D(i)_{jj} = \sum_k \overline{A_{(i)jk}}$. $\overline{A_{(i)}} = \frac{A_{(i)} + A_{(i)}^T}{2}$ is the symmetric affinity matrix of W. The spectral matrix $E_{(i)}$ can be defined using the Laplacian eigenmaps function:

$$\begin{aligned} E_{(i)} &= \min_{E_{(i)}^T E_{(i)} = I} \quad Tr(E_{(i)}^T L E_{(i)}) \\ &= \min_{E_{(i)}^T E_{(i)} = I} \left\| E_{(i)}E_{(i)}^T - A_{(i)} \right\|_F^2. \end{aligned} \tag{17}$$

And the integrated matrix F corresponds to a latent fused affinity matrix S derived from A. We reasonably believe that $S = \mathrm{Diag}(A_{(i)}\mathbf{1}_N) = I_N$. In other words, S can be approximated by its spectral matrix F.

$$F = \min_{F^T F = I} \left\| FF^T - S \right\|_F^2. \tag{18}$$

E. GENERATING BASE CLUSTERS BASED ON DIVERGENCE ENTROPY WITH RPFCM

In order to generate base clusters, we hybridize multiple FCM with RF iterations to generate rich information in different sub-spaces with KL divergence and local diversity. Then, we create low-dimensional data by RP, as stated in Section A. However, we are unable to obtain the optimal q. Therefore, one of the feasible methods is generating multiple low-dimensional data sets simultaneously. After obtaining X_l, we perform FCM on them. Similarly to this, we typically apply multiple FCMs to each low-dimensional data set in turn due to the randomness of the initialization of the FCM.

Another consideration is the number of objects in the ensemble. For the FCM method, the membership matrix U can be thought of as a rich feature set that contains more nuanced information than the clustering indicator matrix [4].

$$Y = [Y_{ij}] \in R^{N\times k}. \tag{19}$$

where $Y_{ij} = \{0, 1\}$ and $y_k^i = 1$. As a result, we use ensemble on $\{U_{(v)}\}_{v=1}^{M}$ rather than $\{Y_{(v)}\}_{v=1}^{M}$.

The fuzzy divergence weight indicator matrix, which is essentially a symmetric matrix, is defined as follows:

$$M_{ij} = \frac{1}{M}\sum_{m=1}^{M} w_i Y_{ij}. \tag{20}$$

$$w_i = DECI(cl(x_i)). \tag{21}$$

where $cl(x_i)$ represents the cluster in $\Pi^m \in \Pi$. Having generated cluster-wise diversity, we further apply spectral learning to obtain final clustering.

F. FUZZY CONSENSUS SPECTRAL LEARNING

In this section, we will introduce how to obtain the final clustering results by fuzzy anchor spectral learning [6].

For the F ensemble clustering, it should incorporate the approximate information of each base partition of all samples via the spectral matrix:

$$\min_{F^TF=I} \quad J = \sum_{i=1}^{M}\left\|FF^T - E_{(i)}E_{(i)}^T\right\|_F^2. \tag{22}$$

To compensate for the loss caused by spectral approximation, we assume that the diagonal matrices $\{\Sigma_i\}_{i=1}^{M}$ can be written as follows:

$$\begin{aligned}\min_{F,\Sigma_i} \quad & J = \sum_{i=1}^{M}\left\|FF^T - E_{(i)}\Sigma_i E_{(i)}^T\right\|_F^2\\ s.t. \quad & FF^T = I;\\ & \left\|\Sigma_i\right\|_F = 1; \quad \Sigma_i \in \Sigma_K;. \end{aligned}\tag{23}$$

Inspired by the anchor graph we introduce the Fuzzy clustering based on anchor graph [6].

$$\begin{aligned}\min_{F,\alpha_i,\Sigma_i} \quad & J = \sum_{i=1}^{M}\alpha_i^m\left\|FF^T - E_{(i)}\Sigma_i E_{(i)}^T\right\|_F^2\\ s.t. \quad & FF^T = I;\\ & \left\|\Sigma_i\right\|_F = 1; \quad \Sigma_i \in \Sigma_K;\\ & \sum_{i=1}^{M}\alpha_i = 1; \quad \alpha_i \succeq 0;. \end{aligned}\tag{24}$$

where α_i is the weight of the base component controlled by the fuzzy exponent m. When $m \to \infty$, the ensemble will usually flatten, and each cluster will have a weight of $1/M$.

Lemma 1 [17]: The multiplicity K of the eigenvalue 0 of the Laplacian matrix L is equal to the number of connected components in the corresponding graph with the similarity matrix.

Inspired by Lemma 1, we build a graph G of a Laplacian matrix L with K zero eigenvalues. The Laplacian matrix L satisfies the following equation: $rank(L) = n - K$. By enforcing that the sum of the L smallest eigenvalues of K tends to 0, the Laplacian rank is convenient to optimize [17].

$$rank(L) = n{-}K \iff \min_{C^TLC} Tr(C^TLC). \tag{25}$$

Finally, by incorporating Equations (14) and (20) as well as (21) into a unified ensemble model. The optimization problem described below leads to the final ensemble clustering:

$$\begin{aligned}\min_{F,\alpha_i,\Sigma_i,S,C} \quad & J = \sum_{i=1}^{M}\alpha_i^m\left\|FF^T - E_{(i)}\Sigma_i E_{(i)}^T\right\|_F^2\\ & + \lambda\left\|FF^T - S\right\|_F^2 + \beta Tr(C^TL_SC)\\ s.t. \quad & FF^T = I; \quad C^TC = I;\\ & \left\|\Sigma_i\right\|_F = 1; \quad \Sigma_i \in \Sigma_K;\\ & \sum_{i=1}^{M}\alpha_i = 1; \quad \alpha_i \succeq 0;\\ & 0 \le S_{ij} \le 1; \quad S_{ij} = 0;. \end{aligned}\tag{26}$$

To solve the problem in Eq. (22), we use an alternative process to optimize F, S, C, α_i, and Σ_i.

1) Fixing F and C, the objective function about S can be expressed as:

$$\begin{aligned}\min_S \quad & \left\|FF^T{-}S\right\|_F^2 + \beta Tr(C^TL_SC)\\ & = \min_S \quad \left\|S{-}QQ\right\|_F^2\\ s.t. \quad & S1_n = 1_n; \quad 0 \le S_{ij} \le 1; \quad S_{ij} = 0;. \end{aligned}\tag{27}$$

where $Tr(C^TL_SC) = \frac{1}{2}\sum_i\sum_i ||x^i - x^i||^2 S_{ij} = Tr(HHS)$ and $QQ = FF^T - \frac{\beta}{4\lambda}HH$

The goal is to find the matrix S that minimizes the objective function. The optimization method is used to minimize the (FF^T) and the prediction matrix (QQ) while satisfying the constraints associated with the affinity matrix.

2) Fixing S and C, the objective function about F can be expressed as:

$$\min_{F,\alpha_i,\Sigma_i,S,C} \quad J = \sum_{i=1}^{M} \alpha_i^m \left\| FF^T - E_{(i)}\Sigma_i E_{(i)}^T \right\|_F^2 + \lambda \left\| FF^T - S \right\|_F^2 \quad s.t. \quad FF^T = I. \tag{28}$$

This objective function is changed to:

$$\frac{\partial \zeta}{\partial F} = \sum_{i=1}^{M} r_i \alpha_i^m \frac{\partial \left\| FF^T - E_{(i)}\Sigma_i E_{(i)}^T \right\|_F^2}{\partial F} + \lambda \frac{\partial \left\| FF^T - S \right\|_F^2}{\partial F} - \partial \frac{\prec \Gamma, FF^T - I \succ}{\partial F}. \tag{29}$$

The goal is to minimize the function associated with F while preserving the constraint $FF^T = I$.

3) Fixing F and S, the objective function about C can be expressed as:

$$min_C \quad \beta Tr(C^T L_S C) \quad s.t. \quad C^T C = I; . \tag{30}$$

where $L_S = \frac{S+S^T}{2}$

The goal is to minimize the workload while satisfying the constraint $C^T C = I$.

We summarize the main steps of our method as follows:

Algorithm 1 Proposed Method

Input: Datasets X; base partition matrices; c is the number of clusters; and the maximum number of iterations $t_{\max}$

Output: Clustering results

1: **Initialization:** Set $t = 0$; Initialize S, F, and ensemble sparse affinity matrix.
2: Run the RPFCM method m times to cluster samples into c clusters.
3: Compute the fuzzy KL divergence entropy of the clusters for further analysis.
4: Generate m fuzzy divergence anchors based on the computed entropy.
5: Convert membership matrix into an anchor graph with KL divergence, forming the ensemble sparse affinity matrix S.
6: **Repeat**
7: Update S via solving Eq. (27) to update the ensemble sparse affinity matrix S.
8: Update F via solving Eq. (28) to update the spectral matrix F.
9: Update C via solving Eq. (30) to update the ensemble sparse affinity matrix C.
10: $t \leftarrow t + 1$.
11: **until** convergence or $t \geq t_{\max}$

IV. EXPERIMENTS RESULTS AND ANALYSIS

In this section, we evaluate the proposed method on 10 high dimensional datasets and compare it with different state-of-the-art ensemble methods.

A. SYNTHETIC DATASETS

In our experiments, 10 image datasets are used, namely, ORL, Yale, YaleB, AR, PIX10, FERET, JAFFE, COIL20, MNIST, and USPS (as shown in TABLE 1). The details of data sets are described as follows:

TABLE 1. Datasets description.

Datasets	Instances	Attributes	Classes
ORL	400	1024	20
Yale	165	1024	15
YaleB	2414	1024	38
AR	2600	2200	100
PIX10	100	54	10000
FERET	1400	1024	200
JAFFE	213	1024	10
COIL20	1440	1024	20
MNIST	5.000	784	10
USPS	9298	256	10

ORL [18] includes 400 human face images, comprising 10 different images, where features have been extracted from 4 views in GIST (512), LBP (59), HOG (864), and CENTRIST (254).

Yale [19] includes 165 images in GIF format of 15 subjects with a variety of expressions and configurations, such as center light, with or without glasses, left-light, right-light, sleepy, surprised, normal, happy, sad, and wink.

Yale B [19] includes 2,414 face images with a size of 192 × 168 over 38 subjects and about 64 images per subject under various illuminations.

AR [20] includes over 4,000 images of 126 subjects. These images include various facial expressions, illumination conditions, and disguises.

PIX10 [21] includes 100 images of 100 × 100 pixels supplied by 10 persons.

FERET [22] includes 1400 images of 200 different people, each with seven images. In the experiment, the size of each image was 80 × 80 pixels.

JAFFE [23] includes 213 images posed by 10 Japanese female models with seven facial expressions (neutral, sadness, happiness, surprise, disgust,anger, and fear).

COIL20 [24] includes 1440 images belonging to 20 different objects. All 72 images were taken from different directions. In the experiment, the size of each image was 32 × 32 pixels.

MNIST [25] consists of 700,000 handwritten digit images with a size of 28 to 28 pixels and is a widely used benchmark dataset.

TABLE 2. The results of comparison experiments.

Datasets		FCM	FCEKL	DLWECDL	PF-CEL	FESRPF	PFREFF	Proposed
ORL	NMI	75.81	77.16	71.71	75.65	80.17	79.83	**83.01**
	ACC	58.46	59.46	54.55	59.7	65.75	**65.99**	65.92
	ARI	62.10	63.09	66.60	66.59	69.70	70.21	**70.23**
Yale	NMI	47.02	48.22	47.12	47.60	52.73	52.99	**53.02**
	ACC	40.59	42.71	45.05	43.09	46.06	**52.99**	53
	ARI	47.50	40.01	47.30	48.30	47.02	48	**49.16**
YaleB	NMI	6.88	6.73	6.90	6.83	7.23	8.21	**8.50**
	ACC	6.61	6.56	6.59	6.57	7	**8**	**8**
	ARI	7	7.19	10	**12.53**	7.50	9.99	12.52
AR	NMI	42.13	42.42	43.62	40.24	43.26	44.19	**44.20**
	ACC	14.58	14.62	14.58	14.70	15.58	**16.33**	16.31
	ARI	15.00	15.15	21.88	20.95	16.15	20.10	**22.81**
PIX10	NMI	87.63	86.99	92.99	93.95	94.4	94.51	**94.56**
	ACC	82.33	80.47	81.02	93.07	94	94.98	**97.42**
	ARI	80	84.09	87.71	84.47	84.64	86	**88.60**
FERET	NMI	48.81	65.92	69.02	61.75	69.02	70.01	**70.98**
	ACC	17.61	24.12	23.79	25.89	29	27	**33.14**
	ARI	61.02	63.47	61.22	63.73	68	66	**69.56**
JAFFE	NMI	89.35	91.58	94.31	94.04	96.04	96.01	**98.99**
	ACC	87.14	88.94	88.90	94.93	97.18	97.32	**97.92**
	ARI	88	90.41	94.11	95.84	97.18	97	**97.98**
COIL20	NMI	73.93	76.12	71.66	73.34	75.6	**77.04**	77.03
	ACC	60.35	64.54	65.31	73.34	75.6	75.99	**76.22**
	ARI	66.01	67.05	73.01	**72.55**	68.53	71.30	71.25
MNIST	NMI	49.46	53.61	50.07	53.94	55.59	56	**58.74**
	ACC	53.38	53.61	53.81	53.94	55.59	55.62	**56.01**
	ARI	55.03	57.06	62.10	63.45	62.71	63	**65.10**
USPS	NMI	59.49	62.81	62.01	61.87	62.36	64.23	**67.99**
	ACC	63.20	67	65.79	67.05	67.96	67.99	**68.11**
	ARI	71.02	73.41	72.09	72.11	73.69	73.98	**73.71**

USPS [26] includes 9298 16-digit handwritten images, which are then divided into 7291 training images and a 2007 test image, which includes all digits 0-9.

Three steps of pre-processing are first applied to the datasets. To create the data matrix X, the images have to be transformed into vectors and straightened. We then convert each vector x_i to a unit vector. The pixel values x_{ij} are then scaled to the [0, 1] range.

B. EVALUATION METRICS

In this paper, three commonly used metrics are used for evaluation: normalized mutual information (NMI), accuracy (ACC) [27], and adjusted Rand index (ARI) [28]. The NMI, ACC, and ARI values are in the range of [0, 1], [0, 1], and $[-1, 1]$, respectively.

C. EXPERIMENT STEP

All the experiments were conducted in MATLAB R2017a on a 64-bit Microsoft Windows 10 (64 bit) computer with an Intel Core i5-2410 M CPU having a memory of 8 GB at 2.30 GHz. In our simulations, we compare the proposed method with state-of-the-art clustering methods, namely, low-rank tensor regularized fuzzy clustering (FCEKL) [29], divergence-based locally weighted ensemble clustering with dictionary learning and L2,1-Norm (DLWECDL) [30], Parameter-Free Robust Ensemble Framework of Fuzzy Clustering (PFREFF) [31], parameter-free consensus embedding learning for multiview graph-based clustering (PF-CEL) [10], and robust and fuzzy ensemble framework via spectral learning for random projection-based fuzzy-c-means clustering (FESRPF) [6].

D. PARAMETER SETTING

The presentation results for the proposed method involve two parts, including base clustering generation and the consensus ensemble. For base clustering, the key parameter is q, which is initially set at 30. The exponent of weighting m is a variable parameter on its ideal interval $(1, 2]$, controlling the contribution ratio of each basis cluster. We adjust it to $m = [1, 1.1, 1.2, 1.3, 1.4, 1.5, 2]$ specifically. To produce the base clustering, the ensemble size M is set to 30. The parameter θ is set to 0.4 to adjust the influence of the cluster uncertainty over the index. We set λ to $[0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1]$ as a coupling parameter that connects F and S.

We use the grid search method to tune the parameters and report the best results for methods in comparison.

E. RESULTS ANALYSIS

The clustering results obtained by different comparison methods on datasets are illustrated in TABLE 2. From the table, we can make the following observations:

1) RP is suitable for FCM because, on most datasets, RPFCM performs as well as or better than FCM.
2) Compared with RPFCM, the proposed method can achieve better results than the best base clustering on most datasets.

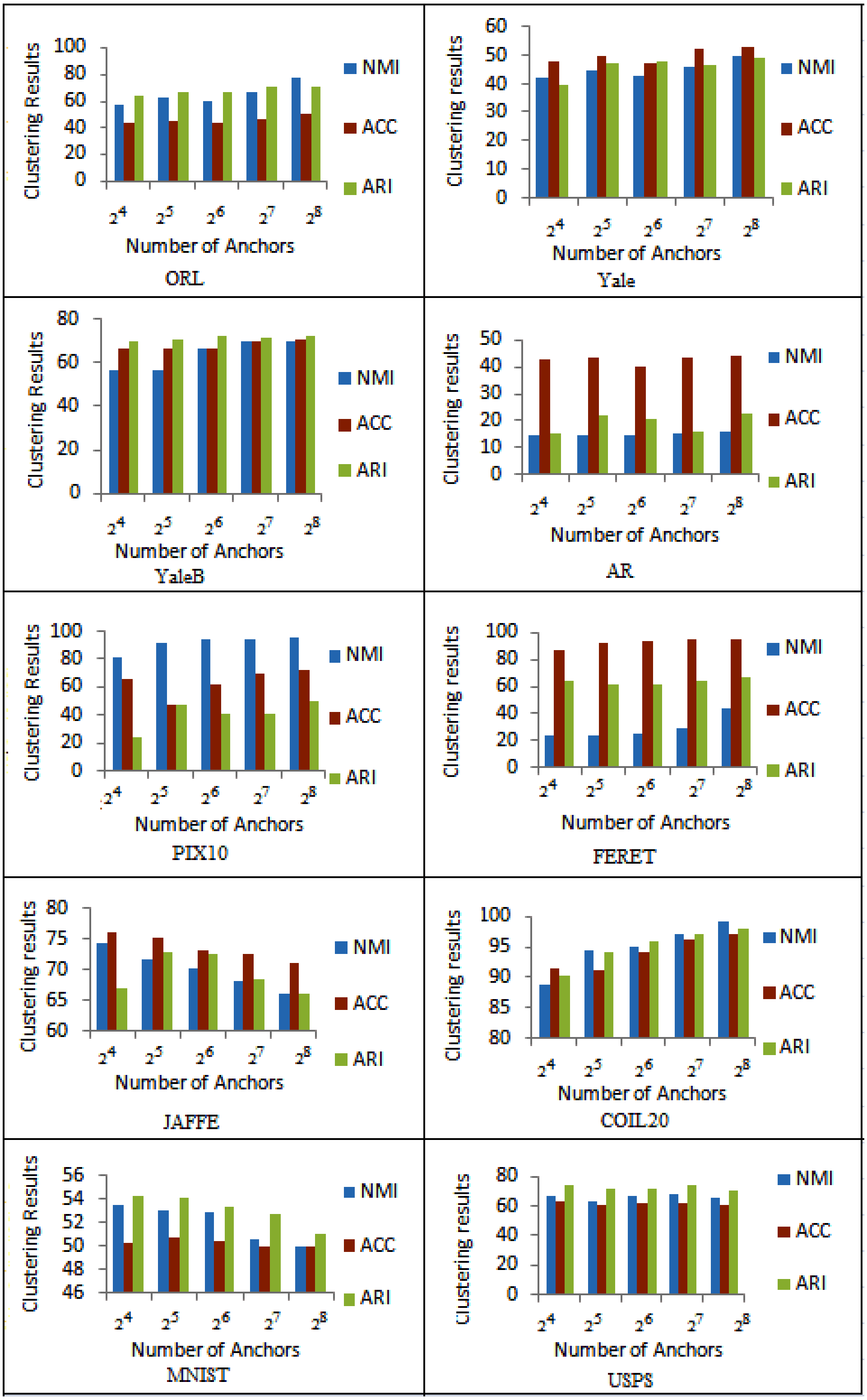


FIGURE 2. **The performance of proposed method with various number of fuzzy anchors.**

3) Ensemble clustering performance is always better than FCM; this can be explained by the fact that ensemble clustering methods are much more accurate than single clustering methods, but they are computationally expensive.

4) We evaluate the proposed method and the other clustering methods using various ensemble sizes M for each run. As can be seen, the proposed method outperforms the other clustering methods in terms of NMI, ACC, and ARI scores. The phenomenon proves

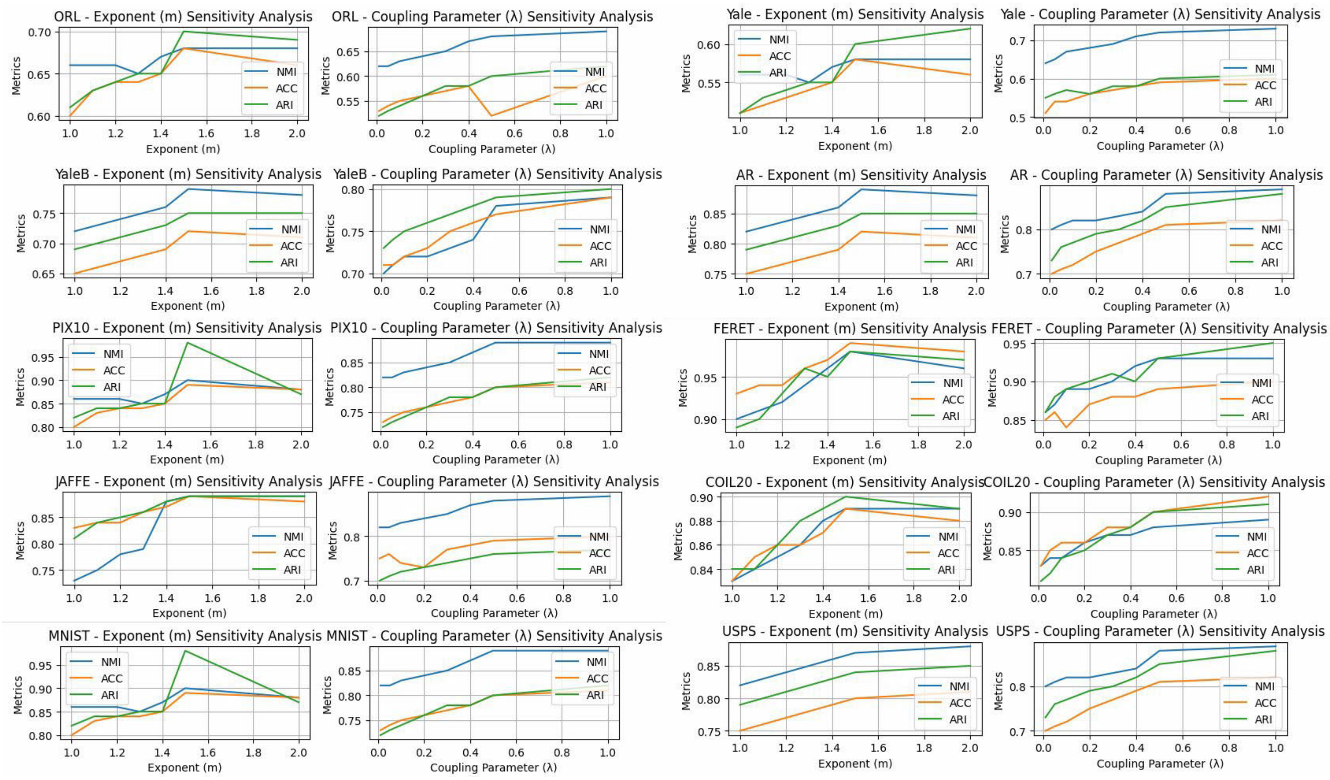


FIGURE 3. Sensitivity NMI, ACC, and ARI about m and λ.

the effectiveness of fusing fuzzy spectral learning with KL divergence and the RP model.

In addition, the proposed method consists of different components, including the use of ensembles, RP, FCM, KL divergence entropy, and spectral learning. Among them, RP is entirely random, and there is no overfitting issue. In each base clustering, the number of clusters in FCM is fixed at the number of real labels. The final clustering results further improve the robustness of the proposed method and prevent the disruption brought on by outlier base clusters.

In FIGURE 2, we present the performance of our proposed clustering method on multiple datasets with varying numbers of fuzzy anchors m in the range $[2^4, 2^5, 2^6, 2^7, 2^8]$. Each line corresponds to a specific evaluation metric (NMI, ACC, ARI) for different anchor numbers. Notably, the trends in clustering performance differ across datasets, emphasizing the impact of anchor selection on overall results. For ORL, Yale, YaleB, AR, PIX10, FERET, and COIL-20 datasets, the number of anchors is set to 28, while for JAFFE, MNIST, and USPS. Furthermore, for the JAFFE, MINIST, and USPS datasets, the clustering performance first rises and then decreases with the increase in anchor numbers. As a result, the number of anchors is set to 28 for JAFFE, 22 for JAFFE, and 27 for USPS, respectively. Analyzing the phenomenon, the main cause is attributed to the quality of the similarity graphs constructed. When the number of anchors increases, the similarity graphs may completely investigate the intricate data structure as the number of anchors rises. However, when the number of anchors rises, it may select more noise, which lowers the quality of similarity graphs. As a result, it is unable to create a similarity graph of good quality.

F. PARAMETER ANALYSIS

To ensure the robustness and effectiveness of the proposed method, we conducted a comprehensive parameter analysis. The key parameters under consideration are $m = [1, 1.1, 1.2, 1.3, 1.4, 1.5, 2]$, the exponent of weighting, and the coupling parameter $\lambda = [0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1]$ (see FIGURE 3).

However, a value around $m = 1.5$ shows slightly better performance in terms of NMI, ACC, and ARI. From FIGURE 3, it can be observed that the sensitivities of parameters are not the same in different data. A decent selection of parameters is important to obtain the best performance.Notably, the sensitivity of these parameters varied across the investigated ranges, underscoring the need for tailored parameter selection for optimal performance.

G. ABLATION EXPERIMENTS

We designed an ablation study to demonstrate the necessity of each component in the proposed method, excluding RP, fuzzy clustering, KL-divergence, and spectral learning, emphasizing their roles in aspects such as dimensionality reduction, cluster quality assessment, convergence, and pattern identification. The results of the ablation experiment are summarized in TABLE 3. From this table, we can

TABLE 3. The results of ablation experiments.

Datasets		RP	Fuzzy Clustering	KL-Divergence	Spectral Learning	Proposed
ORL	NMI	79	83.15	86.14	84.65	**90.20**
	ACC	81.81	83.18	77.55	80.7	**88.99**
	ARI	82.10	79.99	80	80.15	**88.14**
Yale	NMI	55.5	53.22	54.14	60.15	**76.17**
	ACC	54.24	51.12	52	54.51	**72.11**
	ARI	55	53.18	53	55.17	**72.14**
YaleB	NMI	91.28	83.99	91.02	92	**99.10**
	ACC	88.99	83.99	84.15	80.57	**99.21**
	ARI	87.18	88.99	86.2	87.15	**98.55**
AR	NMI	47	42.02	45	48.52	**50.11**
	ACC	81.81	83.18	77.55	80.7	**88.99**
	ARI	46.19	44	45.80	45.12	**51.90**
PIX10	NMI	83.15	83.22	84.44	84.99	**97.20**
	ACC	91.21	91.02	92.05	92.07	**94.98**
	ARI	87.14	87.99	87.25	88.15	**96.56**
FERET	NMI	65.22	65.05	66.82	64.69	**70.15**
	ACC	66.35	66.03	67.55	67.89	**71.02**
	ARI	62.55	68.81	68.84	70.02	**73.25**
JAFF	NMI	88.32	88.15	86.99	88.07	**92.99**
	ACC	88.55	88.18	88.77	89.7	**93**
	ARI	87.58	86.99	88.99	89.15	**92.12**
COIL20	NMI	75.32	74.15	76.25	76.55	**80.23**
	ACC	81.25	80.03	80.98	83.27	**87.21**
	ARI	68.06	67.02	67.80	67.95	**72.25**
MNIST	NMI	59	57.15	57.14	58.65	**60.32**
	ACC	53.81	53.38	52.55	53.27	**55.9**
	ARI	50.10	50.36	52.03	52.15	**54.14**
USPS	NMI	69.36	69.15	68.14	69.39.65	**70.45**
	ACC	68	67.18	67.95	70.27	**72.25**
	ARI	67.99	67.89	68.12	68.69	**70.98**

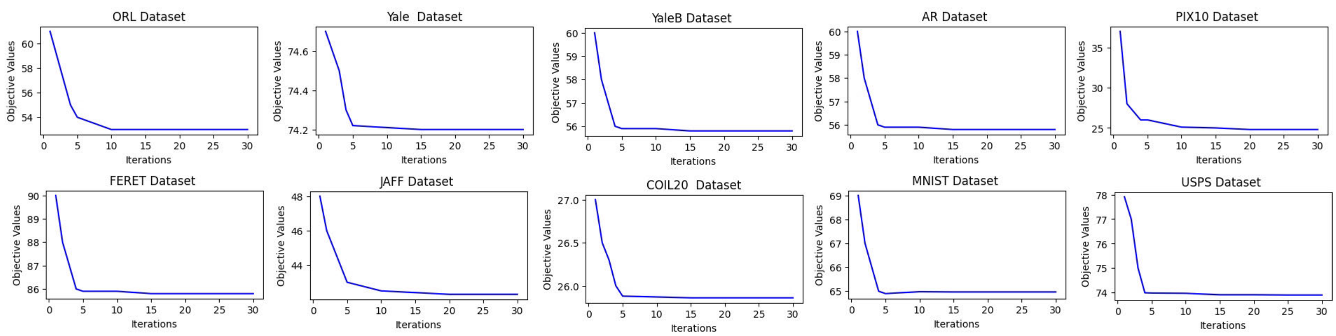


FIGURE 4. The convergence curves.

observe the contribution of each component to the overall effectiveness of the method.

From the results, we can confirm that the data structure extracted from the base set helps our method store the data structure more accurately compared with the individual components across most datasets. The performance of the our method with and without each component removed. In all cases, removing any component leads to a decrease in performance compared to the full method. This proposed that each component plays a crucial role in achieving the optimal clustering results.

H. CONVERGENCE AND EXECUTION TIME

FIGURE 4 shows the convergence curves of 10 datasets. For each data set, the objective value for every iteration is recorded. Obviously, the x-axis represents the number of iterations, and the y-axis represents the target value. From FIGURE 4, it can be seen that the objective value decreases monotonically as the number of iterations increases. Moreover, the proposed method generally converges in less than 30 iterations. These results show that the method is effective.

In addition, we evaluate the execution times of different ensemble clustering methods. To obtain reliable results, the output is averaged over 20 runs. Additionally, it is important to note that as the dimensions and sample numbers increase, the processing costs of clustering methods also tend to increase. However, the proposed method indicates comparable time efficiency compared to the other ensemble clustering methods, as shown in TABLE 4.

I. DISCUSSION

Results demonstrating the superior performance of our fuzzy clustering ensemble method can be explained by a

TABLE 4. The execution times values of different clustering ensemble in seconds.

Data set	FCM	FCEKL	DLWECD	PF-CEL	FESRPF	PFREFF	Proposed
ORL	39.52	**39.11**	44	45	45.02	43	42.19
Yale	**27.2**	39.5	40	51.1	55	45.02	40.03
YaleB	58.6	57.6	**50**	64.6	63.5	61.02	60
AR	90.7	82.2	82.9	70.8	75.1	66	**57.6**
PIX10	46	44.03	43.2	44	45.18	43	**40.06**
FERET	57.1	51	51.7	53	57	48.05	**49.99**
JAFFE	58.5	55	55.8	57	55.98	52.85	**50**
COIL20	81.1	80	80.2	79	79.98	59.02	**60.01**
MNIST	80.7	78	77.09	73	70	68	**66.7**
USPS	87.7	81.15	82.2	77.1	75.98	60.66	**50.7**

number of important factors, including the combination of fuzzy clustering, spectral learning, and RP. It seeks to address challenges with robustness, scalability, and hidden structure finding. To validate its efficacy and investigate its applicability in many fields, more investigation and testing are required. The RP step is used to create multidimensional data; fuzzy partitions of RP are ranked using an entropy-based local weighting along with KL divergence to detect any uncertainty; and spectral learning allows for a more accurate clustering result. Ensemble clustering and fuzzy clustering algorithms, in contrast to classic clustering algorithms, are much more resistant to outliers and noisy data.

The proposed method is superior to FCM clustering in terms of ACC, NMI, and ARI metrics, which may be attributed to ensemble clustering algorithms higher accuracy compared to single clustering algorithms despite their higher computational cost (see TABLE 2).

V. CONCLUSION

In this paper, we present a new fuzzy ensemble clustering method based on spectral learning using RPFCM and KL divergence locally weighted clusters. With the base clustering defined, a fuzzy divergence entropy-based metric is utilized to evaluate and weight the clusters considering the distribution of the cluster labels in the entire ensemble. Furthermore, we convert these membership matrices produced by base clustering into symmetric graphs and approximate the affinity matrices using spectral matrices. Finally, we design an optimization method to jointly learn the modified spectral features and their fusion. The experimental results have been validated on high-dimensional datasets, demonstrating the benefits of the proposed method over other similar ensemble clustering methods. Exploring optimization in ensemble clustering should be an interesting future development.

REFERENCES

[1] M.-S. Yang and Y. Nataliani, "A feature-reduction fuzzy clustering algorithm based on feature-weighted entropy," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 2, pp. 817–835, Apr. 2018.

[2] N. Ilc, "Weighted cluster ensemble based on partition relevance analysis with reduction step," *IEEE Access*, vol. 8, pp. 113720–113736, 2020.

[3] M. Li, W. Liang, and X. Liu, "Multi-view clustering with learned bipartite graph," *IEEE Access*, vol. 9, pp. 87952–87961, 2021.

[4] D. Huang, C.-D. Wang, and J.-H. Lai, "Locally weighted ensemble clustering," *IEEE Trans. Cybern.*, vol. 48, no. 5, pp. 1460–1473, May 2018.

[5] P. Rathore, J. C. Bezdek, S. M. Erfani, S. Rajasegarar, and M. Palaniswami, "Ensemble fuzzy clustering using cumulative aggregation on random projections," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 3, pp. 1510–1524, Jun. 2018.

[6] Z. Shi, L. Chen, J. Duan, G. Chen, and K. Zhao, "Robust and fuzzy ensemble framework via spectral learning for random projection-based fuzzy-c-means clustering," *Eng. Appl. Artif. Intell.*, vol. 117, Jan. 2023, Art. no. 105541.

[7] P. Zhou, L. Du, X. Liu, Y.-D. Shen, M. Fan, and X. Li, "Self-paced clustering ensemble," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 4, pp. 1497–1511, Apr. 2020.

[8] M. Zhang, "Weighted clustering ensemble: A review," *Pattern Recognit.*, vol. 124, Apr. 2022, Art. no. 108428.

[9] D. Wu, Z. Hu, F. Nie, R. Wang, H. Yang, and X. Li, "Multi-view clustering with interactive mechanism," *Neurocomputing*, vol. 449, pp. 378–388, Aug. 2021.

[10] D. Wu, F. Nie, X. Dong, R. Wang, and X. Li, "Parameter-free consensus embedding learning for multiview graph-based clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 7944–7950, Dec. 2022.

[11] P. Rathore, D. Kumar, J. C. Bezdek, S. Rajasegarar, and M. Palaniswami, "A rapid hybrid clustering algorithm for large volumes of high dimensional data," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 4, pp. 641–654, Apr. 2019.

[12] S. K. Sharma, K. Vijayakumar, V. J. Kadam, and S. Williamson, "Breast cancer prediction from MicroRNAs profiling using random subspace ensemble of LDA classifiers via Bayesian optimization," *Multimedia Tools Appl.*, vol. 81, no. 29, pp. 41785–41805, Dec. 2022.

[13] J. Tian, Y. Ren, and X. Cheng, "Stratified feature sampling for semi-supervised ensemble clustering," *IEEE Access*, vol. 7, pp. 128669–128675, 2019.

[14] K. G. Larsen and J. Nelson, "Optimality of the Johnson-lindenstrauss lemma," in *Proc. IEEE 58th Annu. Symp. Found. Comput. Sci. (FOCS)*, Oct. 2017, pp. 633–638.

[15] D. Achlioptas, "Database-friendly random projections," in *Proc. 20th ACM SIGMOD-SIGACT-SIGART Symp. Princ. Database Syst.*, May 2001, pp. 274–281.

[16] W. Liu, J. He, and S. F. Chang., "Large graph construction for scalable semi-supervised learning," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, 2010, pp. 679–686.

[17] F. R. Chung and F. C. Graham, *Spectral Graph Theory*, vol. 92. Providence, RI, USA: Amer. Math. Soc., 1997.

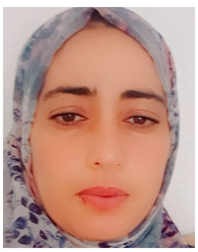
[18] F. S. Samaria and A. C. Harter, "Parameterisation of a stochastic model for human face identification," in *Proc. IEEE Workshop Appl. Comput. Vis.*, Dec. 1994, pp. 138–142.

[19] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001, doi: 10.1109/34.927464.

[20] A. Martinez and R. Benavente, "The AR face database 24," Tech Rep., 1998.

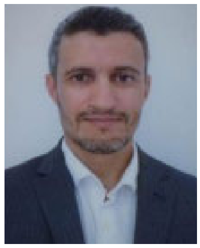
[21] D. Hond and L. Spacek, "Distinctive descriptions for face processing," in *Proc. BMVC*, no. 2, Sep. 1997, p. 4.

[22] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.
[23] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with Gabor wavelets," in *Proc. 3rd IEEE Int. Conf. Autom. Face Gesture Recognit.*, Apr. 1998, pp. 200–205.
[24] A. Sameer, N. Shree K. Nayar, and H. Murase, "Columbia object image library (coil-100)," Tech Rep., 1996.
[25] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
[26] J. Hull, "A database for handwritten text recognition research," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 5, pp. 550–554, May 1994. [Online]. Available: http://dx.doi.org/10.1109/34.291440
[27] A. Strehl and J. Ghosh, "Cluster ensembles: A knowledge reuse framework for combining multiple partitions," *J. Mach. Learn. Res.*, vol. 3, pp. 583–617, Apr. 2003.
[28] N. X. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance," *J. Mach. Learn. Res.*, vol. 11, pp. 2837–2854, 2010.
[29] H. Wei, L. Chen, K. Ruan, L. Li, and L. Chen, "Low-rank tensor regularized fuzzy clustering for multiview data," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 12, pp. 3087–3099, Dec. 2020.
[30] J. Xu, J. Wu, T. Li, and Y. Nan, "Divergence-based locally weighted ensemble clustering with dictionary learning and $L_{2,1}$-norm," *Entropy*, vol. 24, no. 10, p. 1324, Sep. 2022.
[31] Z. Shi, L. Chen, W. Ding, C. Zhang, and Y. Wang, "Parameter-free robust ensemble framework of fuzzy clustering," *IEEE Trans. Fuzzy Syst.*, pp. 1–15, May 2023.



INES LAHMAR received the M.S. degree from the Sfax Institute of Computer Science and Multimedia (ISIMS), University of Sfax, Tunisia, in 2014. Her research interests include big data, fuzzy systems, data mining, and clustering methods.

AIDA ZAIER received the engineering degree from the National Engineering School of Gabes (ENIG), Tunisia, in 2005, and the M.S. degree from the Polytechnic School, Sophia Antipolis, Nice, France, in 2006. Her research interests include channel estimation and synchronization of OFDM and MIMO-OFDM channels under very high mobility conditions, big data, access techniques, and smart technology.



MOHAMED YAHIA received the B.Sc. degree from the Higher School of Sciences and Techniques, Tunis, Tunisia, in 2000, the M.Sc. degree from the Higher School of Communication, Tunis, in 2002, the joint Ph.D. degree in telecommunication engineering from the National Engineering School of Gabes (NESG), Tunisia, and École Nationale Supérieure d'Electrotechnique, d'Electronique, d'Informatique, d'Hydraulique et des Télécommunications (ENSEEIHT), Toulouse, France, in 2010, and the Habilitation degree in telecommunications from the University of Gabes, Tunisia, in 2017. Since 2018, he has been an Associate Professor with NESG. Since 2019, he has been a Research Associate with the American University of Sharjah. His current research interests include image processing, analysis of synthetic aperture radar images, and numerical methods in electromagnetism.

TARIG ALI (Senior Member, IEEE) received the B.S. degree (Hons.) in civil engineering from the University of Khartoum, Sudan, in 1993, and the M.S. and Ph.D. degrees from The Ohio State University, USA, in 1999 and 2003, respectively. He is currently a Professor with the American University of Sharjah. His research interests include geospatial engineering, GIScience, coastal mapping and GIS, and applications of GIS and remote sensing in civil and environmental engineering.

RIDHA BOAULLEGUE was born in Tunis, Tunisia. He received the M.S., Ph.D., and HDR degrees in telecommunications from the National School of Engineering of Tunis (ENIT), Tunisia, in 1990, 1994, and 2003, respectively. He was the Director and the Founder of the National Engineering School of Sousse, in 2005. He was the Director of the School of Technology and Computer Science, in 2010. He is currently the Director of the Innovation of COMmunicant and Cooperative Mobiles Laboratory (INNOV'COM), Sup'COM, Higher School of Communication. His current research interests include mobile and cooperative communications, access techniques, intelligent signal processing, CDMA, MIMO, OFDM, and UWB systems.