Low-Rank Tensor Regularized Graph Fuzzy Learning for Multi-View Data Processing

Baicheng Pan, Chuandong Li, *Senior Member, IEEE,* Hangjun Che, *Member, IEEE,* Man-Fai Leung, *Member, IEEE,* Keping Yu, *Member, IEEE,*

***Abstract*—Multi-view data processing is an effective tool to differentiate the levels of consumers on electronics. Recently, the graph based multi-view clustering methods have attracted widespread attention because they can obtain the relationships of multi-view data points efficiently. However, there exist several shortcomings on most existing graph based clustering methods. Firstly, the mostly adopted Euclidean distance can not extract the nonlinear manifold structure. Secondly, graph based methods are mainly hard clustering methods, which means that each data point belongs to only the one cluster exactly. Thirdly, the high-dimension information between multiple views are not taken into account. Thus, a low-rank tensor regularized graph fuzzy learning (LRTGFL) method for multi-view data processing is proposed. In LRTGFL, Jensen-Shannon divergence is adopted to replace the Euclidean distance for obtaining more completely nonlinear structures. In addition, fuzzy learning is adopted to make graph clustering be a soft clustering method. Furthermore, a tensor nuclear norm based on the tensor singular value decomposition (t-SVD) is adopted to take advantage of the high-dimension information. Then, alternating direction method of multipliers (ADMM) is adopted to solve the LRTGFL model. Finally, the effectiveness and superiority of LRTGFL are demonstrated by comparing with various state-of-the-art algorithms on eight real-world datasets.**

***Index Terms*—Multi-view data processing, graph learning, low-rank tensor, fuzzy clustering.**

I. Introduction

DISTINGUISHING the levels of consumers is of great significance in the field of electronic consumption. By classifying the consumers in different clusters, suppliers can chance services and adjust products with a clear propose. However, the data of consumers which comes from multi-source is inevitably characterized by high-dimensionality, diversity, nonlinearity and complexity. Hence, multi-view clustering is widely applied for multi-view data processing. Besides, it has also attracted considerable attention on internet of things [1], smart home [2], electronic product analysis [3] in field of electronic consumption.

Graph based method is distinguished in clustering due to its simplicity and effectiveness. [4] proposes a graph based clustering method named clustering with adaptive neighbors (CAN). In CAN, the similarity matrix between data points is assigned adaptively based on the local Euclidean distances. The principle of CAN is effective and its structure is simply. Thus CAN is considered as a basic model of graph clustering since it is proposed.

With the popularity of CAN, many multi-view graph clustering methods are proposed. In multi-view datasets, the different views are generated by different ways or origin from a same object. It means that multi-view datasets have more information in variety aspects than single-view datasets. Therefore, it is no doubt that multi-view graph based clustering methods can always obtain more superior results than single-view methods. The most typical and traditional multi-view graph based clustering methods [5–7] obtain a unified graph among the graph similarity matrices of all views first, then some regularization or prior information is added on the unified graph to extract more information and produce the final clusters. However, they have several shortcomings as follows. Firstly, Euclidean distance is always taken to measure the relationships among data points. In Euclidean distance, each data point contributes equally. In fact, the importance of data points at different positions in the sample distribution should be different. In theory, there exist lots of nonlinear manifold structures in many cases on datasets [8]. It is obviously that Euclidean distance can not extract the nonlinear relationships well. Secondly, most existing graph clustering methods are hard clustering methods. Thirdly, high-dimension information between multiple views are not used sufficiently.

In this paper, a low-rank tensor regularized graph fuzzy learning (LRTGFL) method for multi-view data processing is proposed. Firstly, Jensen-Shannon divergence is adopted to represent the relationships between data points instead of traditional Euclidean distance. Jensen-Shannon divergence can measure how much a given arbitrary distribution deviates from the true distribution, and it can find the nonlinear manifold structure between data points and help to obtain a better graph. Secondly, fuzzy clustering is adopted in LRTGFL. Fuzzy clustering can imply the grade of data points which the cluster belongs. Besides, outliers or points which are farther away from the cluster are easy to classify wrongly in hard clustering. Different with hard clustering, outliers can be fuzzy classified in soft clustering, and can be correctly classified after some iterations of computation finally. Thirdly, a tensor singular value decomposition (t-SVD) based tensor nuclear norm is adopted in LRTGFL. Since the number of clusters is much smaller than the number of sample in general, the graph similarity matrices have the block diagonal structure. By

Baicheng Pan, Chuandong Li (*Corresponding author*), and Hangjun Che are with the Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China (e-mail: panbaicheng@email.swu.edu.cn; cdliswu@163.com; hjche123@swu.edu.cn). Man-Fai Leung is with the Faculty of Science and Engineering, Anglia Ruskin University, Cambridge, U.K. (e-mail: man-fai.leung@aru.ac.uk). keping Yu is with the Graduate School of Science and Engineering, Hosei University, Tokyo 184-8584, Japan, and with the RIKEN Center for Advanced Intelligence Project, RIKEN, Tokyo 103-0027, Japan. (e-mail: keping.yu@ieee.org).

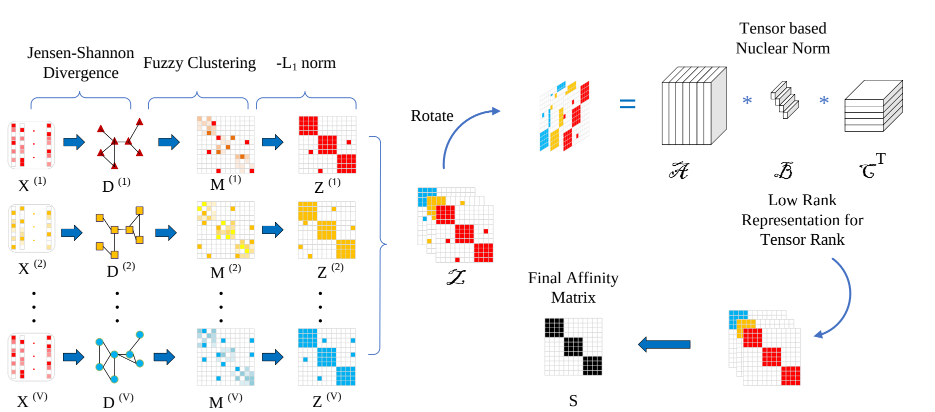


Fig. 1. The flowchart of low-rank tensor regularized graph fuzzy learning (LRTGFL). Firstly, give a multi-view data $X \in \mathbb{R}^{d_v \times N}$. Secondly, compute the Jensen-Shannon divergence between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ and obtain the divergence matrix $\boldsymbol{D}^{(v)}$. Thirdly, make fuzzy clustering to get the membership matrix $\boldsymbol{M}^{(v)}$. Fourthly, $-l_1$ norm is added to membership matrix for avoiding over sparse. Then, construct the similarity tensor $\mathcal{Z}$ by using similarity matrices as its lateral slices. Furthermore, take the t-SVD based tensor nuclear norm to $\mathcal{Z}$. Finally, compute the final affinity matrix by $\boldsymbol{S} = \frac{1}{V}\sum_{v=1}^{v}(\|\boldsymbol{Z}^{(v)}\| + \|\boldsymbol{Z}^{(v)T}\|)$.

constructing the graph similarity tensor from graph similarity matrices, the t-SVD based norm can be a low-rank tensor constraint of it to take high-dimension information between views. The flowchart of the proposed LRTGFL is shown in Fig. 1.

The main contributions of this paper is summarized as follows:

1) A low-rank tensor regularized graph fuzzy learning method for multi-view data processing is proposed in the paper. Nonlinear structures and high-dimension information are taken into account by adopting Jensen-Shannon divergence and low-rank tensor representation respective. Besides, fuzzy clustering is adopted to get a superior clustering performance.
2) An efficient iterative optimization algorithm for LRTGFL based on the alternating direction method of multipliers is presented in the paper.
3) The proposed method is compared with various state-of-the-art methods on eight real-world datasets. According to the clustering performance and the visualization of the experimental results, LRTGFL shows its superiority over the other related clustering methods.

The remaining part of the paper is organized as follows. In Section II, some related works of multi-view clustering are reviewed. In Section III, some notations and preliminaries are introduced briefly. In Section IV, LRTGFL model is proposed and the optimization algorithm is given to solve it. In Section V, the results of experiments which compared with various state-of-the-art methods on eight real-world datasets are shown to verify the superiority of LRTGFL. Finally, the conclusion and outlook of future work are given in Section VI.

II. Related Work

With the development of machine learning, there exist kinds of multi-view clustering methods. They can be roughly classified into four categories: nonnegative matrix factorization based methods, subspace clustering methods, graph based methods and the other categories.

The traditional nonnegative matrix factorization based clustering is a biconvex optimization method [9, 10]. In traditional nonnegative matrix factorization based multi-view clustering method [11, 12], the high-dimension data matrix of each view is decomposed into a basis matrix and a unified representation or coefficient matrix. The structure of nonnegative matrix factorization for clustering is clear, but it is difficult to explore the hidden information. In recent, in order to dig more information in datasets, multiple matrix factorization is proposed. Wang et al. [13] propose a three-matrix factorization based method. [14] proposes the deep matrix factorization which decomposes the data matrix layer-by-layer. According to multiple matrix decomposition, the characteristic hided in deep can be found to get better clustering performance.

Multi-view subspace clustering methods cluster samples into subspaces to find the relationship between data points and clusters. Different with nonnegative matrix factorization, the basis matrix is replaced by a dictionary matrix which is the data matrix itself usually in subspace clustering. By applying the self-representation method, subspace clustering improves its performance. But it is still a significant problem that how to take advantage of the information between views. [15] adopts the Hilbert-Schmidt norm to guarantee the complementary information between views. In fact, information from different views on a object has both consistency and diversity. [16] decomposes the representation matrix into consistency part and specificity part. [17] attempts to harness the consistent and diverse information by introducing a consistency term and a exclusivity term respectively. The tensor approach plays an important role in dealing with high dimensional and complex data [18, 19]. Since exacting the high-dimension information can get a deeper connections between multiple views, the tensor based subspace clustering method is proposed. A tensor based method is proposed by Zhang et al. [20] firstly. However, Zhang's method still adopts the matrix singular value decomposition (SVD) to be a low-rank regularization in fact. Xie et al. [21] introduce a tensor nuclear norm based on tensor singular value decomposition (t-SVD) [22] to multi-view subspace clustering. [23] proposes a multi-view subspace clustering method which takes the low-rank tensor representation and considers the local manifold structure.

Graph based methods has attracted widespread attention due to their clear structure and fast calculation speed. In graph based methods, similarity matrices are constructed by the distance between data to represent the relationship of data points. Some graph based methods [5–7, 24] are working to find a unified graph and construct adaptive weight between multiple views. Auto-weighted algorithms make the clustering work easier and simplify the problem indeed. However, eagerly to construct a unified graph has a shortcoming that it extracts the information among views insufficiently. Therefore, there exists several methods which attempt to extract the connection between multiple views. Wang et al. [25] learn the graph of each view and unified graph simultaneously to get more information. Tang et al. [26] attempt to leverage the complementary information between views by a algorithm named diffusion process. Huang et al. [27] take both the consistency and diversity of graph into account. Besides, high-dimension information is adopted in some works [28–30]. t-SVD based nuclear norm is introduced to add low rank constraint on graph similarity tensor for the high-dimension relationships. In addition, there exists some methods which use projection or nonlinear kernel to replace traditional Euclidean distance for extracting nonlinear structure. Both Wang et al. [31] and Gao et al. [32] project the original data into a low-dimension subspace. Lin et al. [33] embed the original data into a constructed data space nonlinearly. Ren et al. [34] adopt a kind of nonlinear multiple kernel to replace the Euclidean distance.

Except the above kinds of methods, there exists other various multi-view clustering methods, e.g. network based methods [35], support vector machine (SVM) based methods [36], self-supervised learning based methods [37], least squares regression based methods [38], self-paced based methods [39].

III. Notations and Preliminaries

In the section, some notations and preliminaries are introduced briefly. The lower case letters denote the scalars, the bold lowercase letters denote the vectors, the bold capital letters denote matrices, and the bold calligraphy letters denote tensors, e.g., x, $\boldsymbol{x}$, $\boldsymbol{X}$ and $\boldsymbol{\mathcal{X}}$, respectively. $\|\boldsymbol{X}\|_F = \sqrt{\sum_{ij} \boldsymbol{X}_{ij}^2}$ denotes the Frobenius norm of matrix, and $\|\boldsymbol{\mathcal{X}}\|_F = \sqrt{\sum_{ijk} \boldsymbol{\mathcal{X}}_{ijk}^2}$ denotes the Frobenius norm of tensor. $\|\boldsymbol{X}\|_1 = \sum_{ij} |\boldsymbol{X}_{ij}|$ denotes the l_1 norm of matrix. $\|X\|_*$ denotes the matrix nuclear norm based on singular value decomposition, and $\|\boldsymbol{\mathcal{X}}\|_{\circledast}$ denotes the tensor nuclear norm based on tensor singular value decomposition (t-SVD). The Fast Fourier Transformation (FFT) of a tensor $\boldsymbol{\mathcal{X}}$ along the third dimension is denoted by $\overline{\boldsymbol{\mathcal{X}}} = \text{fft}(\boldsymbol{\mathcal{X}}, [\,], 3)$, and the Inverse Fast Fourier Transformation (IFFT) is denoted by $\boldsymbol{\mathcal{X}} = \text{ifft}(\overline{\boldsymbol{\mathcal{X}}}, [\,], 3)$.

Definition 1 (*Tensor Singular Value Decomposition (t-SVD)* [22]). The t-SVD of a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is defined as

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{A}} * \boldsymbol{\mathcal{B}} * \boldsymbol{\mathcal{C}}^T \tag{1}$$

where both $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are orthogonal tensor, $\boldsymbol{\mathcal{C}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a f-diagonal tensor.

Definition 2 (*t-SVD based Tensor Nuclear Norm* [22]). Given a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the value of tensor nuclear norm $\|\boldsymbol{\mathcal{X}}\|_{\circledast}$ is the sum of the singular values on all frontal slices of $\overline{\boldsymbol{\mathcal{X}}}$:

$$\|\boldsymbol{\mathcal{X}}\|_{\circledast} = \sum_{i=1}^{min\{n_1,n_2\}} \sum_{k=1}^{n_3} \|\overline{\boldsymbol{\mathcal{B}}}(i,i,k)\| \tag{2}$$

where $\overline{\boldsymbol{B}}$ is computed by the t-SVD $\overline{\boldsymbol{\mathcal{X}}} = \overline{\boldsymbol{\mathcal{A}}} * \overline{\boldsymbol{\mathcal{B}}} * \overline{\boldsymbol{\mathcal{C}}}^T$.

IV. Methodology

A. Low-Rank Tensor Regularized Graph Fuzzy Learning Model

Divergence is widely applied to measure the relationships and extract nonlinear structures between data points. The Kullback-Leibler (KL) divergence is a popular divergence in machine learning. However, it is an asymmetric divergence. It means that the divergence from $\boldsymbol{x}_i$ to $\boldsymbol{x}_j$ is not equal to the divergence from $\boldsymbol{x}_j$ to $\boldsymbol{x}_i$. In graph clustering, the similarity matrix is hoped to be a symmetric matrix. Therefore, KL divergence is not the best choice to replace the Euclidean distance. Fortunately, the Jensen-Shannon divergence is symmetric to be a better choice. If there exist two distributions $P(i)$ and $Q(i)$, then Jensen-Shannon divergence of them is formulated by

$$D_{JS}(P(i)||Q(i)) = \sum_i P(i)log\frac{P(i)}{M(i)} + Q(i)log\frac{Q(i)}{M(i)}, \tag{3}$$

where $M(i) = (P(i) + Q(i))/2$.

When the divergence between data points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ is large, their similarly z_{ij} should be small. On the contrary, a smaller divergence should be assigned a lager z_{ij}. Fuzzy clustering is a soft clustering, which regards the fuzzy similarity matrix as a suggestive matrix [40]. And it suggests the degree to which points belong to a same cluster. Therefore, a graph based fuzzy clustering method is proposed naturally:

$$\begin{aligned} &\min_{\boldsymbol{Z}^{(v)}} \sum_{v=1}^{V} \sum_{i,j=1}^{N} D_{JS}(\boldsymbol{x}_i^{(v)}||\boldsymbol{x}_j^{(v)}) z_{ij}^{(v)m}, \\ &s.t. \quad z_{ij}^{(v)} \geq 0, \end{aligned} \tag{4}$$

where the V denotes the number of views, N denotes the number of samples, m is the fuzzification factor, $z_{ij}^{(v)}$ is the similarity between $x_i^{(v)}$ and $x_j^{(v)}$.

However, the similarity matrix $\boldsymbol{Z}^{(v)}$ would be over sparse in equation (4) which means that the optimal solution of $\boldsymbol{Z}^{(v)}$ in equation (4) is approaching to $\mathbf{0}$. In order to avoid over sparse of $\boldsymbol{Z}^{(v)}$, The $-l_1$ norm is introduced:

$$\begin{aligned} &\min_{\boldsymbol{Z}^{(v)}} \sum_{v=1}^{V} \sum_{i,j=1}^{N} D_{JS}(\boldsymbol{x}_i^{(v)}||\boldsymbol{x}_j^{(v)}) z_{ij}^{(v)m} - r \sum_{v=1}^{V} \|\boldsymbol{Z}^{(v)}\|_1, \\ &s.t. \quad z_{ij}^{(v)} \geq 0, \end{aligned} \tag{5}$$

where r is a balance parameter of the $-l_1$ norm.

As the multi-view features are extracted from the same objects, there exists consistency between views. Besides, the number of clusters is much smaller than the number of samples [21, 41]. Therefore the t-SVD based nuclear norm which uses similarity matrices as the lateral slices represents low-rank structure can help to obtain the high-dimension information between views. By adding the t-SVD based nuclear norm, the formula is as follows:

$$\min_{\boldsymbol{Z}^{(v)}} \sum_{v=1}^{V} \sum_{i,j=1}^{N} D_{JS}(\boldsymbol{x}_i \| \boldsymbol{x}_j) z_{ij}^{(v)m} - r \sum_{v=1}^{V} \|\boldsymbol{Z}^{(v)}\|_1 + \lambda \|\boldsymbol{\mathcal{Z}}\|_{\circledast}, \quad s.t. \quad z_{ij}^{(v)} \geq 0, \tag{6}$$

where $\|\boldsymbol{\mathcal{Z}}\|_{\circledast}$ is the tensor based nuclear norm, λ is a balance parameter of it.

In order to balance the weight of divergence between data points, a column normalization constraint is added to $\boldsymbol{Z}^{(v)}$. Thus the final objective function is formulated as follows:

$$\min_{\boldsymbol{Z}^{(v)}} \sum_{v=1}^{V} \sum_{i,j=1}^{N} D_{JS}(\boldsymbol{x}_i^{(v)} \| \boldsymbol{x}_j^{(v)}) z_{ij}^{(v)m} - r \sum_{v=1}^{V} \|\boldsymbol{Z}^{(v)}\|_1 + \lambda \|\boldsymbol{\mathcal{Z}}\|_{\circledast}, \quad s.t. \sum_{i=1}^{N} z_{ij}^{(v)} = 1, z_{ij}^{(v)} \geq 0, \tag{7}$$

where $D_{JS}(\boldsymbol{x}_i^{(v)} \| \boldsymbol{x}_j^{(v)})$ denotes the Jansen-Shannon divergence between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ in v-th view, $\boldsymbol{Z}^{(v)}$ is the similarity graph matrix of v-th view, $\boldsymbol{\mathcal{Z}}$ represents the similarity tensor constructed by similarity graph matrices of all views, m means the fuzzification factor, and both r and λ are balance parameter.

B. Optimization

The nuclear norm $\|\boldsymbol{\mathcal{Z}}\|_{\circledast}$ in the objective function (7) is difficult to solve directly. Thus, an auxiliary tensor $\boldsymbol{\mathcal{G}}$ is introduced to solve the problem conveniently. Moreover, the fuzzification factor is fixed as $m = 2$ to simplify the optimization process:

$$\min_{\boldsymbol{Z}^{(v)}, \boldsymbol{\mathcal{G}}} \sum_{v=1}^{V} \sum_{i,j=1}^{N} D_{JS}(\boldsymbol{x}_i^{(v)} \| \boldsymbol{x}_j^{(v)}) z_{ij}^{(v)2} - r \sum_{v=1}^{V} \|\boldsymbol{Z}^{(v)}\|_1 + \lambda \|\boldsymbol{\mathcal{G}}\|_{\circledast} \quad s.t. \sum_{i=1}^{N} z_{ij}^{(v)} = 1, z_{ij}^{(v)} \geq 0, \boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{G}}. \tag{8}$$

Then, the augmented Lagrangian formulation of the problem (8) is expressed as:

$$\min_{\boldsymbol{Z}^{(v)}, \boldsymbol{\mathcal{G}}} \sum_{v=1}^{V} \sum_{i,j=1}^{N} D_{JS}(\boldsymbol{x}_i^{(v)} \| \boldsymbol{x}_j^{(v)}) z_{ij}^{(v)2} - r \sum_{v=1}^{V} \|\boldsymbol{Z}^{(v)}\|_1 + \lambda \|\boldsymbol{\mathcal{G}}\|_{\circledast} + \frac{\mu}{2} \|\boldsymbol{\mathcal{Z}} - \boldsymbol{\mathcal{G}} + \frac{\boldsymbol{\Phi}}{\mu}\|_F^2 + \sum_{v=1}^{V} \sum_{j=1}^{N} (\eta_j^{(v)} (\boldsymbol{z}_j^{(v)T} \mathbf{1} - 1) - \boldsymbol{\beta}_j^{(v)} \boldsymbol{z}_j^{(v)}) \tag{9}$$

where $\boldsymbol{\Phi}$ is a tensor, $\eta_j \geq 0$ is a constant, $\boldsymbol{\beta}_j \geq \mathbf{0}$ is a column vector, they are all Lagrangian multipliers in the algorithm and μ is the penalty parameter and satisfies $\mu \geq 0$.

Then, according to the alternating direction method of multipliers (ADMM) [42], the variables are updated as follows:

Update $\boldsymbol{Z}^{(v)}$. With $\boldsymbol{\mathcal{G}}$, $\boldsymbol{\Phi}$ and μ fixed, $\boldsymbol{Z}^{(v)}$ is updated by

$$\min_{\boldsymbol{Z}^{(v)}} \sum_{v=1}^{V} \sum_{i,j=1}^{N} D_{JS}(\boldsymbol{x}_i^{(v)} \| \boldsymbol{x}_j^{(v)}) z_{ij}^{(v)2} - r \sum_{v=1}^{V} \|\boldsymbol{Z}^{(v)}\|_1 + \lambda \|\boldsymbol{\mathcal{G}}\|_{\circledast} + \frac{\mu}{2} \|\boldsymbol{\mathcal{Z}} - \boldsymbol{\mathcal{G}} + \frac{\boldsymbol{\Phi}}{\mu}\|_F^2 + \sum_{v=1}^{V} \sum_{j=1}^{N} (\eta_j^{(v)} (\boldsymbol{z}_j^{(v)T} \mathbf{1} - 1) - \boldsymbol{\beta}_j^{(v)} \boldsymbol{z}_j^{(v)}) \tag{10}$$

The formula (10) can be solved by the closed-form solution proposed in [4]. Moreover, it can be also solved as follows. The problem (10) can be separated into N subproblems on each view. Thus, the j-th subproblem is:

$$\boldsymbol{z}_j^{(v)*} = \arg\min_{\boldsymbol{z}_j^{(v)}} \sum_{i=1}^{N} d_{ij}^{(v)} \boldsymbol{z}_{ij}^{(v)2} - r \sum_{i=1}^{N} z_{ij}^{(v)} + \frac{\mu}{2} z_{ij}^{(v)2} - \mu z_{ij}^{(v)} (g_{ij}^{(v)} - \phi_{ij}^{(v)}/2) + \eta_j^{(v)} (\boldsymbol{z}_j^{(v)T} \mathbf{1} - 1) - \boldsymbol{\beta}_j^{(v)} \boldsymbol{z}_j^{(v)} \tag{11}$$

where the $d_{ij}^{(v)} = D_{JS}(\boldsymbol{x}_i^{(v)} \| \boldsymbol{x}_j^{(v)})$. Then the vector form of Eq. (11) can be written as:

$$\begin{aligned} \boldsymbol{z}_j^{(v)*} &= \arg\min_{\boldsymbol{z}_j^{(v)}} \|\boldsymbol{z}_j^{(v)} - \frac{\mu(\boldsymbol{g}_j^{(v)} - \boldsymbol{\Phi}_j^{(v)}/2) + r}{\mu + 2\boldsymbol{d}_j^{(v)}}\|_2^2 \\ &\quad + \eta_j^{(v)} (\boldsymbol{z}_j^{(v)T} \mathbf{1} - 1) - \boldsymbol{\beta}_j^{(v)} \boldsymbol{z}_j^{(v)} \\ &= \arg\min_{\boldsymbol{z}_j^{(v)}} \|\boldsymbol{z}_j^{(v)} - \boldsymbol{h}_j^{(v)}\|_2^2 \\ &\quad + \eta_j^{(v)} (\boldsymbol{z}_j^{(v)T} \mathbf{1} - 1) - \boldsymbol{\beta}_j^{(v)} \boldsymbol{z}_j^{(v)}, \end{aligned} \tag{12}$$

where $\boldsymbol{h}_j = \frac{\mu(\boldsymbol{g}_j^{(v)} - \boldsymbol{\Phi}_j^{(v)}/2) + r}{\mu + 2\boldsymbol{d}_j^{(v)}}$. Based on the KKT conditions, we have:

$$\begin{cases} 2z_{ij}^{(v)*} - 2h_{ij}^{(v)*} - \eta_j^{(v)*} - \beta_{ij}^{(v)*} = 0, \\ \beta_{ij}^{(v)*} z_{ij}^{(v)*} = 0, \\ \beta_{ij}^{(v)*} \geq 0 \\ z_{ij}^{(v)*} \geq 0, \end{cases} \tag{13}$$

According to $\boldsymbol{s}_j^{(v)T} \mathbf{1} = 1$, we get

$$\eta_j^{(v)*} = \frac{2 - 2\boldsymbol{h}_j^{(v)*T} \mathbf{1} - \boldsymbol{\beta}_j^{(v)*T} \mathbf{1}}{N} \tag{14}$$

Then according to Eq. (14) and the complementary slackness condition in Eq. (13), the solution of $z_{ij}^{(v)*}$ is:

$$z_{ij}^{(v)*} = max \left\{ h_{ij} + \frac{1 - \boldsymbol{h}_j^{(v)*T} \mathbf{1} - (\boldsymbol{\beta}_j^{(v)*T} \mathbf{1})/2}{N}, 0 \right\}. \tag{15}$$

Update $\boldsymbol{\mathcal{G}}$. With $\boldsymbol{Z}^{(v)}$, $\boldsymbol{\Phi}$ and μ fixed, $\boldsymbol{\mathcal{G}}$ is updated by solving

$$\mathcal{G}^* = \arg\min_{\mathcal{G}} \lambda\|\mathcal{G}\|_{\circledast} + \frac{\mu}{2}\|\mathcal{Z} - \mathcal{G} + \frac{\boldsymbol{\Phi}}{2}\|_F^2 \tag{16}$$

The calculation of optimal solution for problem (16) is in Fourier domain. $\overline{\mathcal{G}}^*$ is update by tensor tubal-shrinkage operator [21]:

$$\overline{\mathcal{G}}^* = \overline{\mathcal{A}} * \mathcal{J}_{\lambda/\mu}\overline{\mathcal{B}} * \overline{\mathcal{C}}^T, \tag{17}$$

where $\overline{\mathcal{B}}$ is calculated by $\overline{\mathcal{T}} = \overline{\mathcal{A}} * \overline{\mathcal{B}} * \overline{\mathcal{C}}^T$, and $\mathcal{T} = \mathcal{Z} + \frac{\boldsymbol{\Phi}}{2}$. In addition, $\mathcal{J}_{\lambda/\mu}\overline{\mathcal{B}}(i,i,j) = max\{\overline{\mathcal{B}}(i,i,j) - \frac{\lambda}{\mu}, 0\}$.

Update $\boldsymbol{\Phi}$ and μ. With $\boldsymbol{Z}^{(v)}$ and $\mathcal{G}$ fixed, $\boldsymbol{\Phi}$ and μ is updated by

$$\begin{cases} \boldsymbol{\Phi}^* = \boldsymbol{\Phi} + \mu(\mathcal{Z} - \mathcal{G}), \\ \mu^* = min\{\gamma * \mu, \mu_{max}\}, \end{cases} \tag{18}$$

where γ satisfies $\gamma > 1$. It is set as $\gamma = 2$ in the paper.

The optimization flowchart of the proposed LRTGFL is summarized in Algorithm 1. The convergence condition is set as follows:

$$max \left\{ \begin{array}{l} \|\boldsymbol{Z}_k^{(v)} - \boldsymbol{G}_k^{(v)}\|_\infty, \\ \|\boldsymbol{Z}_k^{(v)} - \boldsymbol{Z}_{k-1}^{(v)}\|_\infty, \\ \|\boldsymbol{G}_k^{(v)} - \boldsymbol{G}_{k-1}^{(v)}\|_\infty, \end{array} \right\} \le \epsilon, \tag{19}$$

where ϵ is a small positive parameter, and it is set as 10^{-7} in the paper.

Algorithm 1 LRTGFL for multi-view clustering

Input: Multi-view data $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \cdots, \boldsymbol{X}^{(V)}$, parameter r, λ, and the initial value of $\boldsymbol{Z}_0^{(v)}$

Initialize: $\mu_0 = 10^{-3}$, $\gamma = 2$, $\mu_{max} = 10^8$, $\epsilon = 10^{-7}$, $\mathcal{G}_0 = \mathcal{Z}_0$, $\boldsymbol{\Phi}_0 = \boldsymbol{0}$, and set the iterative number $k = 1$;
While not converged **do**
Update $\boldsymbol{Z}^{(v)}$ by Eq. (15);
Update $\mathcal{G}$ by Eq. (17);
Update $\boldsymbol{\Phi}, \mu$ by Eq. (18);
Check the convergence condition in Eq. (19);
$k = k + 1$;
end while
Compute the affinity matrix $\boldsymbol{S} = \frac{1}{V}\sum_{v=1}^{v}(\|\boldsymbol{Z}^{(v)}\| + \|\boldsymbol{Z}^{(v)T}\|)$;

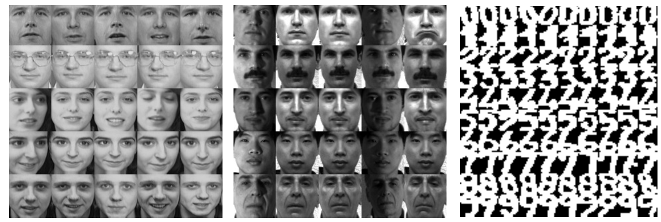
Output: Affinity matrix $\boldsymbol{S}$

C. Complexity Analysis

The complexity is consumed on updating $\boldsymbol{Z}^{(v)}$ and $\mathcal{G}$ mainly. The size of $\boldsymbol{Z}^{(v)}$ is $N \times N$, and the size of $\mathcal{G}$ is $N \times N \times V$. To update $\boldsymbol{Z}^{(v)}$, we need to solve VN sub-problems. To update each $z_{ij}^{(v)}$, we need to calculate $\boldsymbol{h}_j^{(v)*T}\boldsymbol{1}$ and $\boldsymbol{\beta}_j^{(v)*T}\boldsymbol{1}$ in Eq. (15) which costs $\mathcal{O}(N)$. Thus we need $\mathcal{O}(VN^2)$ to update $\boldsymbol{Z}^{(v)}$ in total. To update $\mathcal{G}$, the FFT and IFFT need to be calculated, which takes $\mathcal{O}(VN^2log(N))$. In Fourier domain, the SVD of each frontal slice of $\mathcal{T}$ with size $N \times N \times V$ need to be calculated, which takes $\mathcal{O}(V^2N^2)$. Thus, the overall complexity is $\mathcal{O}(VN^2(1 + log(N) + N))$. In general, there exists $1 < log(N) \ll N$, thus the overall complexity is $\mathcal{O}(VN^3)$ in an iteration.

V. Experiments

A. Datasets



(a) ORL (b) Yale (c) UCI

Fig. 2. Samples of datasets. (a) ORL. (b) Yale. (c) UCI.

Eight common real-world multi-view datasets are adopted in experiments. Their information are summarized on Table I. Some samples of them are shown in Fig. 2. They are introduced as follows briefly:

ORL[1][43]: It is a face images dataset which includes 400 images of 40 different people with three views. Each category of images are collected under different conditions.

Yale[2]: It is also a face images dataset with three views. It includes 165 face images of 15 individuals under different lighting condition and expression on face.

UCI[3]: It is a classical handwritten digits dataset. It includes 2000 samples of 10 classes digits from 0 to 9 with three views.

Handwritten: It is a handwritten digits dataset of 0 to 9 which contains 2000 samples in total with six views.

BBCsport and **BBC4view**[4]: Both BBCsport and BBC4view are datasets which come from BBC news website on 5 types of topic. BBCsport contains 544 documents in total and has two views. BBC4view includes 685 documents and has four views.

NGs[5]: NGs is a news dataset with 500 data samples with three views which categorized into five clusters.

100leaves[6]: It includes 1600 samples of 100 kinds of plant leaves. Three views are extracted including 64 dimensions shape descriptor, 64 dimensions texture histogram and 64 dimensions fine scale margin.

B. Evaluation Metrics

The performance of clustering result is evaluated by six popular evaluation metrics: accuracy (ACC), normalized mutual information (NMI), adjusted rand index (AR), F-score, Precision and Recall. For all metrics, higher value denotes more predominant clustering performance.

C. Comparison Methods

Fourteen clustering methods are taken to compare with the proposed LRTGFL, including three single-view clustering

[1]http://www.uk.research.att.com/facedatabase.html
[2]https://cvc.yale.edu/projects/yalefaces/yalefaces.html
[3]https://archive.ics.uci.edu/ml/datasets/Multiple+Features
[4]http://mlg.ucd.ie/datasets
[5]http://lig-membres.imag.fr/grimal/data.html
[6]https://archive.ics.uci.edu/ml/datasets/One-hundred+plant+species+leaves+data+set

TABLE I
INFORMATION OF REAL-WORLD MULTI-VIEW DATASETS.

Datasets	Objective	Samples	Clusters	View1	View2	View3	View4	View5	Views6
ORL	Face images	400	40	$4096d$	$3304d$	$6750d$	-	-	-
Yale	Face images	165	15	$4096d$	$3304d$	$6750d$	-	-	-
UCI	Handwritten digit	2000	10	$240d$	$76d$	$6d$	-	-	-
Handwritten	Handwritten digit	2000	10	$216d$	$76d$	$64d$	$6d$	$240d$	$47d$
BBCsport	News text	544	5	$3183d$	$3203d$	-	-	-	-
BBC4view	News text	685	5	$4659d$	$4633d$	$4665d$	$4684d$	-	-
NGs	News text	500	5	$2000d$	$2000d$	$2000d$	-	-	-
100leaves	Object	1600	100	$64d$	$64d$	$64d$	-	-	-

methods (SSC[7] [44], LRR[8] [45], CAN [4]), a nonnegative matrix factorization based method (MVCC[9] [13]), four subspace multi-view clustering methods (RMSC [46], LT-MSC [20], MLRSSC[10] [47], CSMSC[11] [16]) and six graph based multi-view clustering methods (MVGL[12] [24], GSF[13] [48], GMC[14] [25], CGD[15] [26], CGL[16] [29], CDMGC[17] [27]):

D. Performance Comparison

The clustering results on eight real world datasets are shown in Table II and III. The experiments are running 10 times, and the average and standard deviation are computing as the performance results. Bold values and underlined values represent the best and second-best results respectively.

On all of the eight datasets, LRTGFL performs best of the six evaluation metrics. On ORL, CGL, UCI, 100leaves and Handwritten, CGL performs second-best. On BBCsport, CGD performs the second-best. On BBC4view, MLRSSC performs the second-best. On NGs, LT-MSC performs second-best. There are 10%, 15%, 4%, 5%, 6%, 12%, 1%, 2% improvement for LRTGFL compared with the second-best performance results on the ORL, Yale, UCI, Handwritten, BBCsport, BBC4view, NGs and 100leaves respectively.

In general, single-view methods can not achieve the performance of multi-view clustering methods. It is obviously that multi-view methods can obtain more information from multiple views.

All of LT-MSC, CGL and the proposed LRTGFL are tensor based clustering methods. LT-MSC still use the SVD nuclear norm based on matrix, but both CGL and LRTGFL take the tensor based SVD nuclear norm and get the high-dimensional information from all views. Therefore, CGL and LRTGFL perform better than LT-MSC in total.

MVGL, GSF, GMC, CGD, CGL, CDMGC and the proposed LRTGFL are all graph based methods. Both MVGL and GSF are committed to find a global graph. GMC, CGD and CDMSC take some consistent or divergent information between views into account. Both CGL and LRTGFL adopt a tensor nuclear norm which can use high-dimension information between views. Therefore, CGL and LRTGFL performs better than the others. In addition, LRTGFL adopts Jensen-Shannon divergence and fuzzy learning which can extract the nonlinear structure in data. Thus it can obtain the connections between data points effectively and correctly.

E. Comparison of Visualization

The visualization of the final affinity matrix generated by similarity matrix or representation matrix is shown in Fig. 3. RMSC and CSMSC are subspace methods, and all of the others are graph based methods. In Fig. 3, white color denotes large values, and black color denotes smaller values. The final affinity matrix of LRTGFL is clearer and the diagonal block structures are completer compared with the others. It indicates that there are more correct connections of graph in LRTGFL.

In Fig. 4, visualization of the clustering performance on UCI is shown. By using the function t-SNE on Matlab, the visualization of the unified graph of each method is generated. Since both CGL and LRTGFL have low-rank tensor regularization on similarity graph tensor, they perform more superior than the others. Besides, it can be seen that LRTGFL reveals the clearest structure of each cluster and has the lowest misclasssfications. Furthermore, because of the strategy of fuzzy clustering, the distance between different clusters is the farthest and the distance between data points in same clusters is lowest of LRTGFL. Thus, LRTGFL is the most superior clustering method of them.

F. Parameter Analysis

For the fuzzification parameter m, we set $m = 2$ in all experiments. In the experiments of LRTGFL, r and λ need to be tuned. The grid search method is adopted for tuning the parameters, and both r and λ are selected from the sets of $[0.005, 0.01, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10]$ in the experiments. ACC values and NMI values of LRTGFL on eight real world datasets with different parameter settings is shown in Fig. 5 and Fig. 6 respectively. Moreover, ORL, Yale, UCI and 100leaves are images datasets and both BBCsport and NGs are text datasets. From Fig. 5 and Fig. 6, it is easy to know that the parameters are less sensitive in images datasets than text datasets.

[7] http://www.ccis.neu.edu/home/eelhami/codes.htm
[8] https://sites.google.com/site/guangcanliu/
[9] https://github.com/vast-wang/Clustering
[10] https://github.com/mbrbic/Multi-view-LRSSC
[11] https://github.com/XIAOCHUN-CAS/Consistent-and-Specific-Multi-View-Subspace-Clustering
[12] https://github.com/kunzhan/MVGL
[13] https://github.com/dugzzuli/Graph-structure-fusion-for-multiview-clustering
[14] https://github.com/cshaowang/gmc
[15] https://github.com/ChangTang/CGD
[16] https://github.com/guanyuezhen/CGL
[17] https://github.com/huangsd/CDMGC

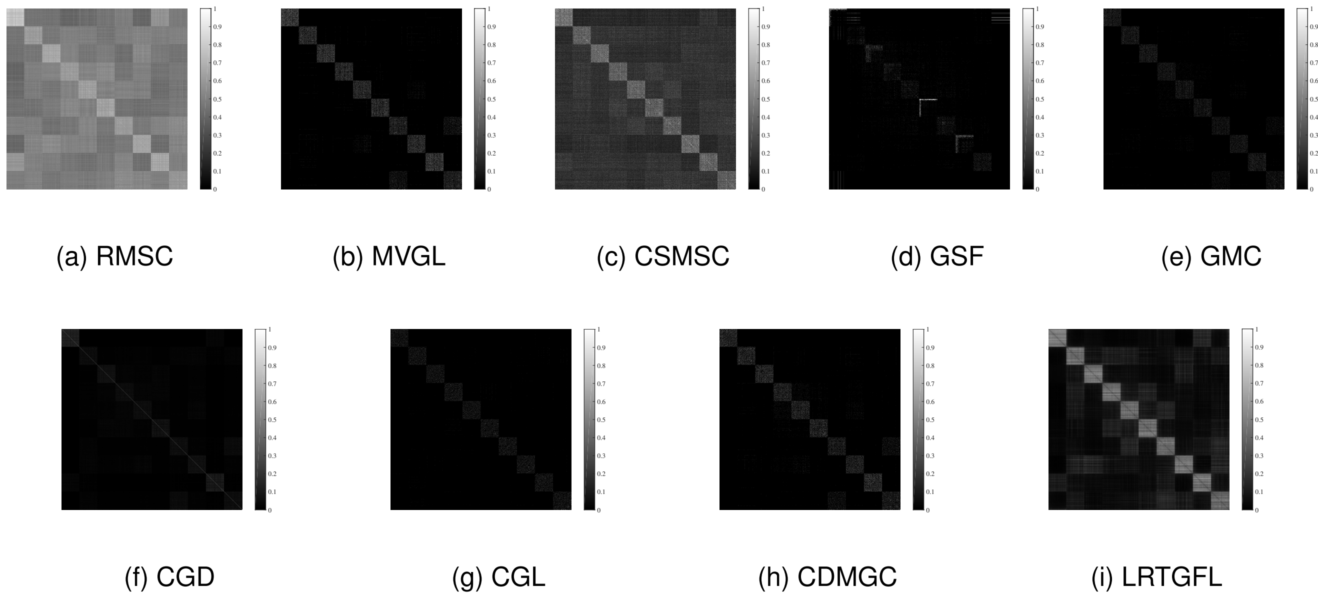


Fig. 3. Visualization of the final affinity matrix on UCI of different methods. (a) RMSC. (b) MVGL. (c) CSMSC. (d) GSF. (e) GMC. (f) CGD. (g) CGL. (h) CDMGC. (i) LRTGFL.

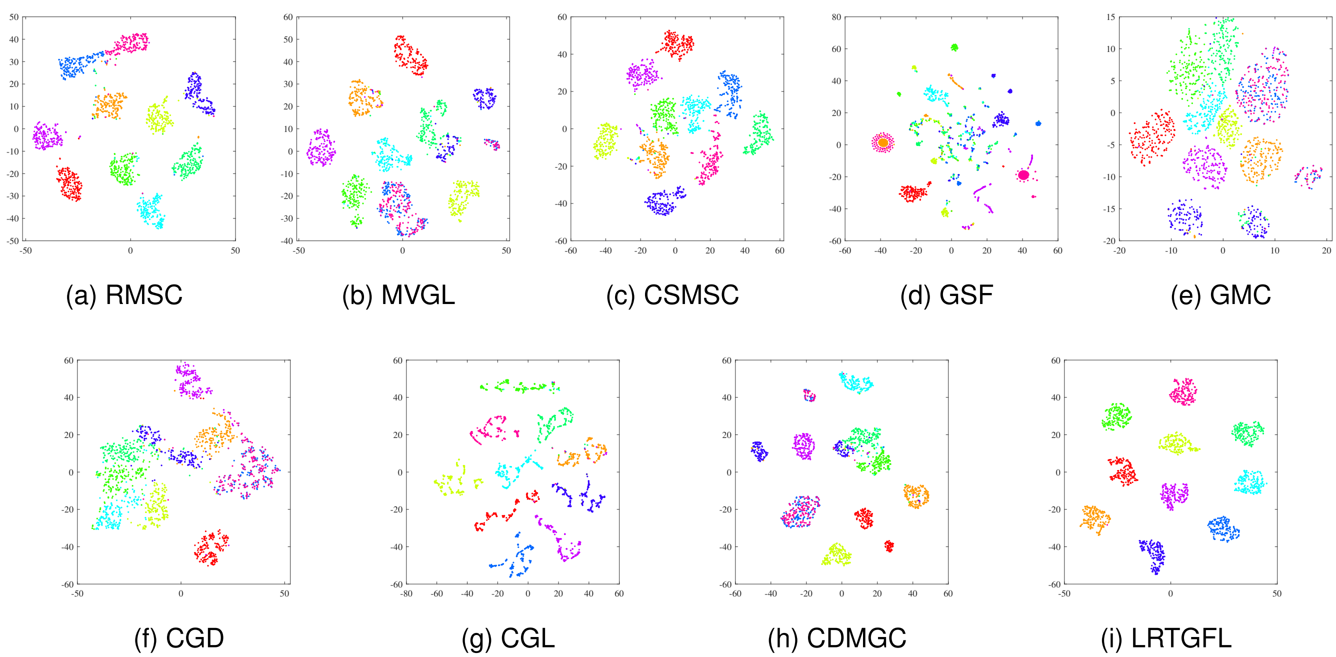
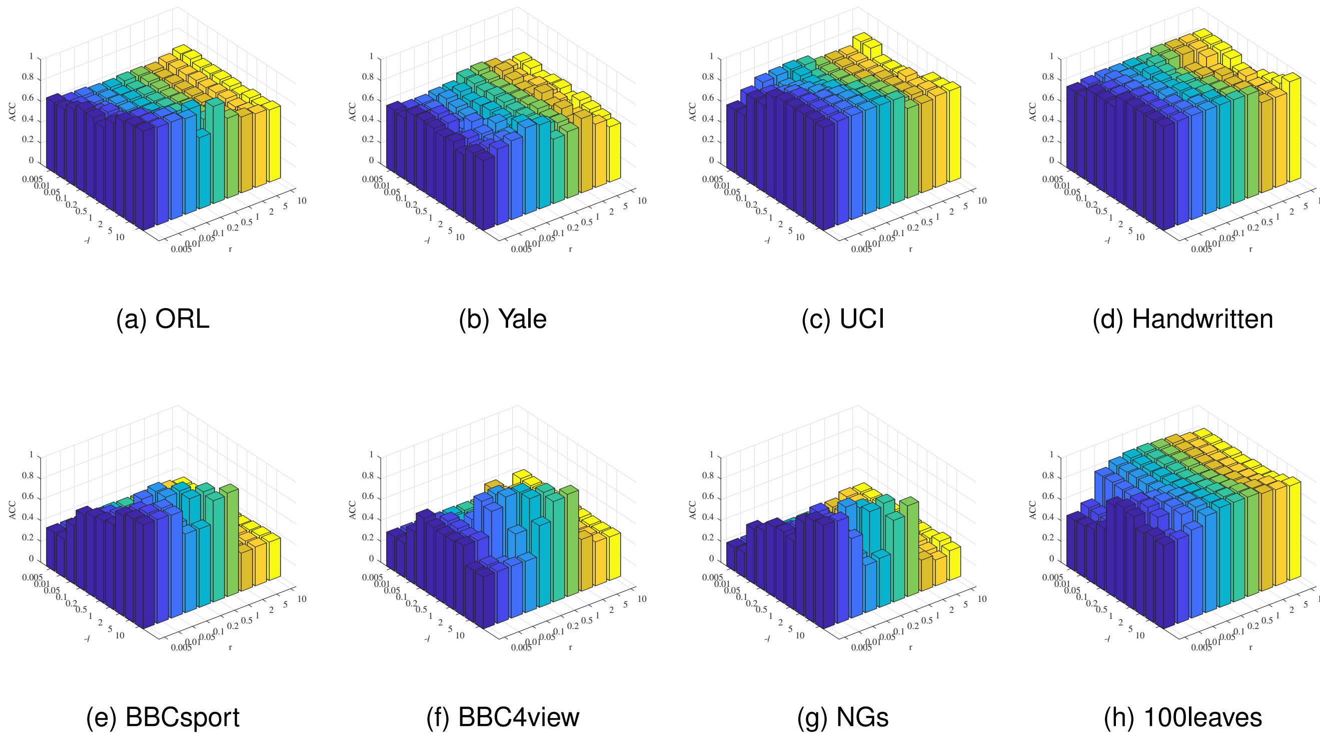


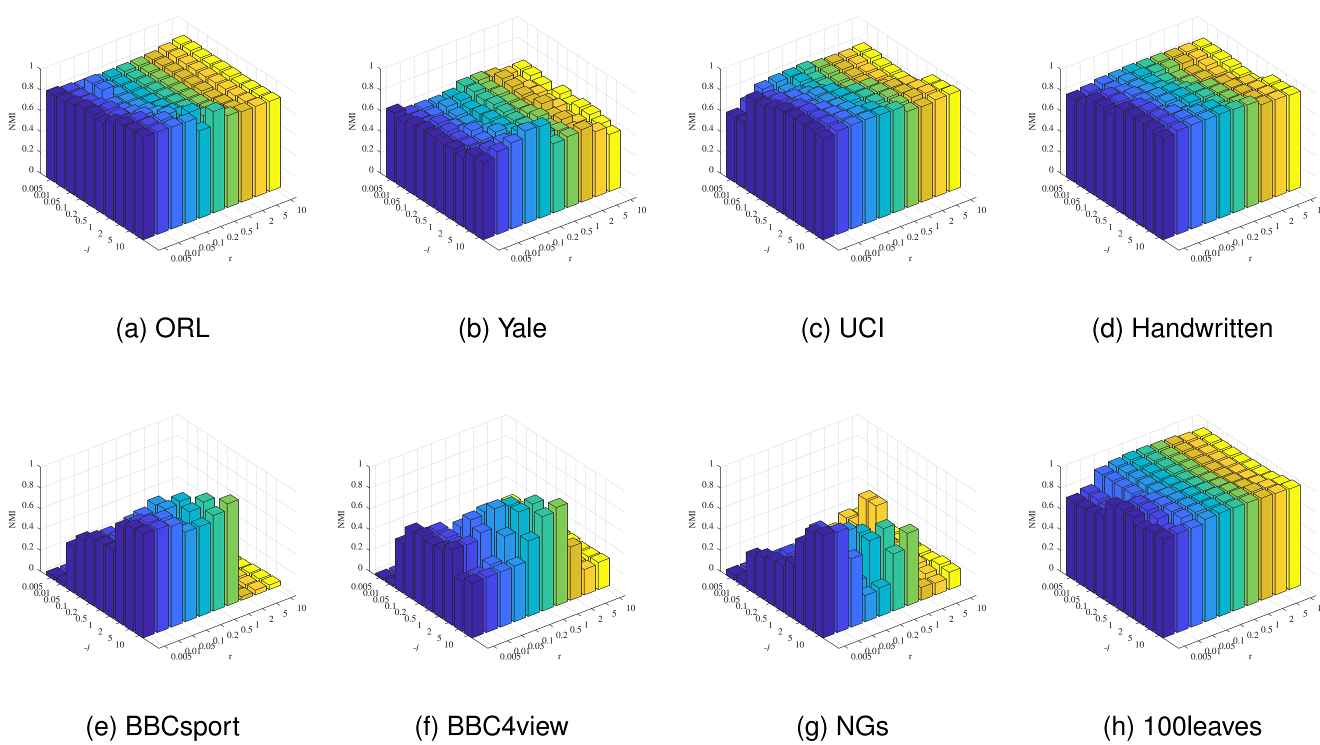
Fig. 4. Visualization of the final affinity matrix on UCI of different methods. (a) RMSC. (b) MVGL. (c) CSMSC. (d) GSF. (e) GMC. (f) CGD. (g) CGL. (h) CDMGC. (i) LRTGFL.



(a) ORL (b) Yale (c) UCI (d) Handwritten

(e) BBCsport (f) BBC4view (g) NGs (h) 100leaves

Fig. 5. ACC values of the proposed method on real world datasets with different parameter settings. (a) ORL. (b) Yale. (c) UCI. (d) Handwritten. (e) BBCsport. (f) BBC4view. (g) NGs. (h) 100leaves.



(a) ORL (b) Yale (c) UCI (d) Handwritten

(e) BBCsport (f) BBC4view (g) NGs (h) 100leaves

Fig. 6. ACC values of the proposed method on real world datasets with different parameter settings. (a) ORL. (b) Yale. (c) UCI. Handwritten. (e) BBCsport. (f) BBC4view. (g) NGs. (h) 100leaves.

TABLE II

CLUSTERING RESULTS (MEAN±STANDARD DEVIATION).

Datasets	Method	ACC	NMI	AR	F-score	Precision	Recall
ORL	SSC	0.6783 ± 0.0147	0.8150 ± 0.0065	0.5430 ± 0.0116	0.5539 ± 0.0113	0.5231 ± 0.0113	0.5887 ± 0.0132
	LRR	0.7023 ± 0.0123	0.8336 ± 0.0073	0.5849 ± 0.0187	0.5949 ± 0.0182	0.5577 ± 0.0212	0.6376 ± 0.0177
	CAN	0.5575 ± 0.0000	0.7628 ± 0.0000	0.2131 ± 0.0000	0.2418 ± 0.0000	0.1496 ± 0.0000	0.6306 ± 0.0000
	RMSC	0.7385 ± 0.0130	0.8686 ± 0.0066	0.6501 ± 0.0146	0.6585 ± 0.0143	0.6205 ± 0.0168	0.7017 ± 0.0144
	LT-MSC	0.8068 ± 0.0211	0.9170 ± 0.0077	0.7530 ± 0.0233	0.7589 ± 0.0227	0.7170 ± 0.0280	0.8063 ± 0.0190
	MVCC	0.3225 ± 0.0000	0.5981 ± 0.0000	0.1417 ± 0.0000	0.1746 ± 0.0000	0.1027 ± 0.0000	0.5806 ± 0.0000
	MVGL	0.6000 ± 0.0000	0.8117 ± 0.0000	0.2529 ± 0.0000	0.2803 ± 0.0000	0.1725 ± 0.0000	0.7467 ± 0.0000
	MLRSSC	0.6948 ± 0.0212	0.8552 ± 0.0086	0.6005 ± 0.0218	0.6104 ± 0.0212	0.5548 ± 0.0254	0.6790 ± 0.0231
	CSMSC	0.7867 ± 0.0114	0.9029 ± 0.0048	0.7208 ± 0.0136	0.7275 ± 0.0132	0.6834 ± 0.0196	0.7781 ± 0.0111
	GSF	0.8000 ± 0.0000	0.9074 ± 0.0000	0.7149 ± 0.0000	0.7219 ± 0.0000	0.6639 ± 0.0000	0.7911 ± 0.0000
	GMC	0.6765 ± 0.0024	0.8697 ± 0.0023	0.4814 ± 0.0120	0.4972 ± 0.0115	0.3693 ± 0.0127	0.7611 ± 0.0038
	CGD	0.6900 ± 0.0000	0.8341 ± 0.0000	0.5138 ± 0.0000	0.5278 ± 0.0000	0.4131 ± 0.0000	0.7306 ± 0.0000
	CGL	$\underline{0.8670} \pm \underline{0.0093}$	$\underline{0.9284} \pm \underline{0.0026}$	$\underline{0.8106} \pm \underline{0.0086}$	$\underline{0.8150} \pm \underline{0.0084}$	$\underline{0.7974} \pm \underline{0.0113}$	$\underline{0.8334} \pm \underline{0.0063}$
	CDMGC	0.6450 ± 0.0000	0.8388 ± 0.0000	0.2693 ± 0.0000	0.2961 ± 0.0000	0.1821 ± 0.0000	0.7928 ± 0.0000
	LRTGFL	**0.9408 ± 0.0189**	**0.9849 ± 0.0049**	**0.9389 ± 0.0185**	**0.9403 ± 0.0181**	**0.9148 ± 0.0251**	**0.9675 ± 0.0112**
Yale	SSC	0.5909 ± 0.0135	0.6467 ± 0.0105	0.4035 ± 0.0149	0.4413 ± 0.0141	0.4240 ± 0.0124	0.4602 ± 0.0171
	LRR	0.5145 ± 0.0101	0.5433 ± 0.0110	0.3015 ± 0.0125	0.3425 ± 0.0117	0.3367 ± 0.0118	0.3524 ± 0.0116
	CAN	0.4909 ± 0.0000	0.5486 ± 0.0000	0.2331 ± 0.0000	0.2959 ± 0.0000	0.2202 ± 0.0000	0.4509 ± 0.0000
	RMSC	0.5491 ± 0.0152	0.5936 ± 0.0058	0.2622 ± 0.0157	0.3198 ± 0.0125	0.2505 ± 0.0196	0.4453 ± 0.0171
	LT-MSC	0.7352 ± 0.0029	0.7636 ± 0.0065	0.5895 ± 0.0154	0.6156 ± 0.0141	0.5914 ± 0.0203	0.6421 ± 0.0081
	MVCC	0.4182 ± 0.0000	0.4614 ± 0.0000	0.1512 ± 0.0000	0.2276 ± 0.0000	0.1542 ± 0.0000	0.4339 ± 0.0000
	MLRSSC	0.5394 ± 0.0000	0.3753 ± 0.0000	0.3389 ± 0.0000	0.4206 ± 0.0000	0.6069 ± 0.0000	0.3301 ± 0.0000
	CSMSC	0.6473 ± 0.0206	0.6957 ± 0.0055	0.4923 ± 0.0073	0.5252 ± 0.0069	0.4943 ± 0.0057	0.5604 ± 0.0095
	GSF	0.6242 ± 0.0000	0.6964 ± 0.0000	0.4856 ± 0.0000	0.5188 ± 0.0000	0.4898 ± 0.0000	0.5515 ± 0.0000
	GMC	0.6645 ± 0.0000	0.6892 ± 0.0000	0.4410 ± 0.0000	0.4801 ± 0.0000	0.4188 ± 0.0000	0.5624 ± 0.0000
	CGD	0.6121 ± 0.0000	0.6553 ± 0.0000	0.4490 ± 0.0000	0.4846 ± 0.0000	0.4566 ± 0.0000	0.5164 ± 0.0000
	CGL	$\underline{0.7515} \pm \underline{0.0000}$	$\underline{0.7737} \pm \underline{0.0000}$	$\underline{0.6254} \pm \underline{0.0017}$	$\underline{0.6488} \pm \underline{0.0016}$	$\underline{0.6336} \pm \underline{0.0010}$	$\underline{0.6647} \pm \underline{0.0023}$
	CDMGC	0.7091 ± 0.0000	0.7399 ± 0.0000	0.3950 ± 0.0000	0.4406 ± 0.0000	0.3560 ± 0.0000	0.5782 ± 0.0000
	LRTGFL	**0.8582 ± 0.0544**	**0.9068 ± 0.0234**	**0.8043 ± 0.0567**	**0.8167 ± 0.0530**	**0.7913 ± 0.0644**	**0.8445 ± 0.0434**
UCI	SSC	0.7423 ± 0.0000	0.7175 ± 0.0000	0.6313 ± 0.0000	0.6619 ± 0.0000	0.6496 ± 0.0000	0.6898 ± 0.0000
	LRR	0.6792 ± 0.0025	0.7480 ± 0.0042	0.6354 ± 0.0039	0.6757 ± 0.0038	0.6084 ± 0.0017	0.7599 ± 0.0119
	CAN	0.8290 ± 0.0000	0.8931 ± 0.0000	0.8157 ± 0.0000	0.8353 ± 0.0000	0.7811 ± 0.0000	0.8976 ± 0.0000
	RMSC	0.9337 ± 0.0000	0.8676 ± 0.0000	0.8587 ± 0.0000	0.8278 ± 0.0000	0.8713 ± 0.0000	0.8743 ± 0.0000
	LT-MSC	0.7985 ± 0.0086	0.7699 ± 0.0093	0.7177 ± 0.0131	0.7463 ± 0.0118	0.7333 ± 0.0111	0.7598 ± 0.0125
	MVCC	0.4450 ± 0.0000	0.5022 ± 0.0000	0.3010 ± 0.0000	0.3751 ± 0.0000	0.3519 ± 0.0000	0.4016 ± 0.0000
	MVGL	0.8400 ± 0.0000	0.8633 ± 0.0000	0.7826 ± 0.0000	0.8057 ± 0.0000	0.7558 ± 0.0000	0.8626 ± 0.0000
	MLRSSC	0.8180 ± 0.0000	0.7748 ± 0.0000	0.7243 ± 0.0000	0.8327 ± 0.0000	0.8229 ± 0.0000	0.7479 ± 0.0000
	CSMSC	0.8821 ± 0.0016	0.8041 ± 0.0013	0.7650 ± 0.0026	0.7886 ± 0.0023	0.7830 ± 0.0025	0.7942 ± 0.0021
	GSF	0.5365 ± 0.0000	0.6032 ± 0.0000	0.2871 ± 0.0000	0.3905 ± 0.0000	0.2680 ± 0.0000	0.7193 ± 0.0000
	GMC	0.6170 ± 0.0000	0.7688 ± 0.0000	0.5604 ± 0.0000	0.6133 ± 0.0000	0.5075 ± 0.0000	0.7853 ± 0.0000
	CGD	0.8250 ± 0.0000	0.8212 ± 0.0000	0.7418 ± 0.0000	0.7680 ± 0.0000	0.7515 ± 0.0000	0.7853 ± 0.0000
	CGL	$\underline{0.9688} \pm \underline{0.0447}$	$\underline{0.9537} \pm \underline{0.0138}$	$\underline{0.9445} \pm \underline{0.0422}$	$\underline{0.9502} \pm \underline{0.0376}$	$\underline{0.9432} \pm \underline{0.0571}$	$\underline{0.9582} \pm \underline{0.0146}$
	CDMGC	0.7085 ± 0.0000	0.8029 ± 0.0000	0.6827 ± 0.0000	0.6716 ± 0.0000	0.5783 ± 0.0000	0.8808 ± 0.0000
	LRTGFL	**0.9980 ± 0.0000**	**0.9948 ± 0.0000**	**0.9956 ± 0.0000**	**0.9960 ± 0.0000**	**0.9960 ± 0.0000**	**0.9960 ± 0.0000**
Handwritten	LRR	0.5483 ± 0.0001	0.5824 ± 0.0000	0.4333 ± 0.0010	0.4916 ± 0.0001	0.4752 ± 0.0001	0.5093 ± 0.0010
	CAN	0.7245 ± 0.0000	0.7711 ± 0.0000	0.6415 ± 0.0000	0.6820 ± 0.0000	0.6020 ± 0.0000	0.7865 ± 0.0000
	RMSC	0.9048 ± 0.0001	0.8230 ± 0.0016	0.8039 ± 0.0018	0.8235 ± 0.0016	0.8197 ± 0.0017	0.8273 ± 0.0015
	LT-MSC	0.9021 ± 0.0142	0.8384 ± 0.0120	0.8060 ± 0.0214	0.8254 ± 0.0792	0.8006 ± 0.0199	0.8281 ± 0.0186
	MVGL	0.8610 ± 0.0000	0.9097 ± 0.0000	0.8381 ± 0.0000	0.8554 ± 0.0000	0.7935 ± 0.0000	0.9297 ± 0.0000
	MLRSSC	0.7005 ± 0.0000	0.6566 ± 0.0000	0.6140 ± 0.0000	0.7075 ± 0.0000	0.7145 ± 0.0000	0.6157 ± 0.0000
	CSMSC	0.9100 ± 0.0001	0.8422 ± 0.0000	0.9149 ± 0.0001	0.8335 ± 0.0001	0.8297 ± 0.0001	0.8372 ± 0.0001
	GMC	0.8820 ± 0.0000	0.9050 ± 0.0000	0.8502 ± 0.0000	0.8658 ± 0.0000	0.8264 ± 0.0000	0.9093 ± 0.0000
	CGD	0.8560 ± 0.0000	0.8911 ± 0.0000	0.8354 ± 0.0000	0.8530 ± 0.0000	0.7916 ± 0.0000	0.9247 ± 0.0000
	CGL	$\underline{0.9770} \pm \underline{0.0000}$	$\underline{0.9495} \pm \underline{0.0000}$	$\underline{0.9496} \pm \underline{0.0000}$	$\underline{0.9546} \pm \underline{0.0000}$	$\underline{0.9540} \pm \underline{0.0000}$	$\underline{0.9552} \pm \underline{0.0000}$
	CDMSC	0.8460 ± 0.0000	0.9103 ± 0.0000	0.8318 ± 0.0000	0.8498 ± 0.0000	0.7931 ± 0.0000	0.9152 ± 0.0000
	LRTGFL	**1.0000 ± 0.0000**	**1.0000 ± 0.0000**	**1.0000 ± 0.0000**	**1.0000 ± 0.0000**	**1.0000 ± 0.0000**	**1.0000 ± 0.0000**

On Table II and Table III, the parameters of the proposed LRTGFL are set as follows. We set $r = 0.01$ and $\lambda = 1$ in ORL, $r = 0.1$ and $\lambda = 10$ in Yale, $r = 0.05$ and $\lambda = 5$ in UCI, $r =$ and $\lambda =$ in Handwritten, $r = 0.005$ and $\lambda = 10$ in BBCsport, $r =$ and $\lambda =$ in BBC4view, $r = 0.05$ and $\lambda = 1$ in NGs, $r = 0.5$ and $\lambda = 5$ in 100leaves, respectively.

Table IV and Table V show the p-values between the proposed LRTGFL and other clustering methods. LRTGFL significantly outperforms the most of the others on most of the datasets.

G. Ablation Analysis

In this section, the effectiveness of different components in LRTGFL is verified by ablation experiments, i.e., the Jensen-Shannon divergence, t-SVD based nuclear norm and fuzzification.

a)The effectiveness of the Jensen-Shannon divergence

In order to verify the effectiveness of the Jensen-Shannon divergence, the Jensen-Shannon divergence is replaced by the Euclidean distance in this case. For convenience, this ablate method is called as low-rank tensor regularized graph fuzzy learning with Euclidean distance (LRTGFL-ED). The objective function of LRTGFL-ED is as follows:

TABLE III
CLUSTERING RESULTS (MEAN±STANDARD DEVIATION).

Datesets	Method	ACC	NMI	AR	F-score	Precision	Recall
	SSC	0.7346 ± 0.0016	0.4830 ± 0.0017	0.4364 ± 0.0026	0.5777 ± 0.0018	0.5496 ± 0.0025	0.6087 ± 0.0010
	LRR	0.8915 ± 0.0000	0.8023 ± 0.0000	0.8328 ± 0.0000	0.8730 ± 0.0000	0.8667 ± 0.0000	0.8793 ± 0.0000
	CAN	0.7004 ± 0.0000	0.6479 ± 0.0000	0.4532 ± 0.0000	0.6261 ± 0.0000	0.4724 ± 0.0000	0.9280 ± 0.0000
	RMSC	0.8952 ± 0.0000	0.8264 ± 0.0000	0.8414 ± 0.0000	0.8797 ± 0.0000	0.8706 ± 0.0000	0.8889 ± 0.0000
	LT-MSC	0.5656 ± 0.0365	0.3290 ± 0.0289	0.2588 ± 0.0736	0.5007 ± 0.0344	0.3724 ± 0.0593	0.7901 ± 0.0599
	MVCC	0.8585 ± 0.0000	0.8127 ± 0.0000	0.7931 ± 0.0000	0.8462 ± 0.0000	0.7863 ± 0.0000	0.9160 ± 0.0000
	MVGL	0.9688 ± 0.0000	0.9044 ± 0.0000	0.9155 ± 0.0000	0.9358 ± 0.0000	0.9294 ± 0.0000	0.9423 ± 0.0000
BBCsport	MLRSSC	0.8143 ± 0.0000	0.7361 ± 0.0000	0.7511 ± 0.0000	0.8116 ± 0.0000	0.7963 ± 0.0000	0.8275 ± 0.0000
	CSMSC	0.9485 ± 0.0000	0.8490 ± 0.0000	0.8645 ± 0.0000	0.8970 ± 0.0000	0.8927 ± 0.0000	0.9014 ± 0.0000
	GSF	0.7574 ± 0.0000	0.7946 ± 0.0000	0.6131 ± 0.0000	0.7272 ± 0.0000	0.5880 ± 0.0000	0.9258 ± 0.0000
	GMC	0.7390 ± 0.0000	0.7954 ± 0.0000	0.6099 ± 0.0000	0.7207 ± 0.0000	0.5728 ± 0.0000	0.9714 ± 0.0000
	CGD	0.9743 ± 0.0000	0.9126 ± 0.0000	0.9305 ± 0.0000	0.9472 ± 0.0000	0.9421 ± 0.0000	0.9523 ± 0.0000
	CGL	0.9449 ± 0.0000	0.8553 ± 0.0000	0.8597 ± 0.0000	0.8919 ± 0.0000	0.9286 ± 0.0000	0.8579 ± 0.0000
	CDMGC	0.7371 ± 0.0000	0.7911 ± 0.0000	0.5969 ± 0.0000	0.7187 ± 0.0000	0.5697 ± 0.0000	0.9712 ± 0.0000
	LRTGFL	**1.0000 ± 0.0000**	**1.0000 ± 0.0000**	**1.0000 ± 0.0000**	**1.0000 ± 0.0000**	**1.0000 ± 0.0000**	**1.0000 ± 0.0000**
	SSC	0.6467 ± 0.0000	0.3678 ± 0.0000	0.3733 ± 0.0000	0.5240 ± 0.0000	0.4722 ± 0.0000	0.6361 ± 0.0000
	LRR	0.7088 ± 0.0016	0.4945 ± 0.0024	0.4606 ± 0.0017	0.5926 ± 0.0013	0.5681 ± 0.0011	0.6193 ± 0.0016
	CAN	0.4175 ± 0.0000	0.2284 ± 0.0000	0.0787 ± 0.0000	0.4134 ± 0.0000	0.2669 ± 0.0000	0.9163 ± 0.0000
	RMSC	0.7299 ± 0.0000	0.5275 ± 0.0000	0.4646 ± 0.0000	0.5897 ± 0.0000	0.5918 ± 0.0000	0.5876 ± 0.0000
	LT-MSC	0.7270 ± 0.0088	0.5633 ± 0.0053	0.5668 ± 0.0047	0.6654 ± 0.0039	0.6564 ± 0.0027	0.6464 ± 0.0049
	MVCC	0.7620 ± 0.0000	0.5996 ± 0.0000	0.6277 ± 0.0000	0.7175 ± 0.0000	0.6978 ± 0.0000	0.7383 ± 0.0000
	MVGL	0.7007 ± 0.0000	0.5852 ± 0.0000	0.5019 ± 0.0000	0.6488 ± 0.0000	0.5159 ± 0.0000	0.8741 ± 0.0000
BBC4view	MLRSSC	0.9328 ± 0.0000	0.8857 ± 0.0000	0.8341 ± 0.0000	0.8874 ± 0.0000	0.8093 ± 0.0000	0.8506 ± 0.0000
	CSMSC	0.8788 ± 0.0000	0.7047 ± 0.0000	0.7303 ± 0.0000	0.7935 ± 0.0000	0.7946 ± 0.0000	0.7923 ± 0.0000
	GSF	0.3883 ± 0.0000	0.2641 ± 0.0000	0.1219 ± 0.0000	0.3678 ± 0.0000	0.3081 ± 0.0000	0.4563 ± 0.0000
	GMC	0.6905 ± 0.0000	0.5536 ± 0.0000	0.4748 ± 0.0000	0.6301 ± 0.0000	0.4999 ± 0.0000	0.8519 ± 0.0000
	CGD	0.8861 ± 0.0000	0.7227 ± 0.0000	0.7688 ± 0.0000	0.8268 ± 0.0000	0.7730 ± 0.0000	0.8887 ± 0.0000
	CGL	0.8628 ± 0.0000	0.7071 ± 0.0000	0.7134 ± 0.0000	0.7773 ± 0.0000	0.8178 ± 0.0000	0.7407 ± 0.0000
	CDMGC	0.4818 ± 0.0000	0.3788 ± 0.0000	0.1572 ± 0.0000	0.4539 ± 0.0000	0.3024 ± 0.0000	0.9092 ± 0.0000
	LRTGFL	**0.9912 ± 0.0000**	**0.9676 ± 0.0000**	**0.9798 ± 0.0000**	**0.9845 ± 0.0000**	**0.9890 ± 0.0000**	**0.9801 ± 0.0000**
	SSC	0.5840 ± 0.0000	0.4699 ± 0.0000	0.3513 ± 0.0000	0.5064 ± 0.0000	0.4201 ± 0.0000	0.6372 ± 0.0000
	LRR	0.5240 ± 0.0000	0.3671 ± 0.0000	0.2256 ± 0.0000	0.4205 ± 0.0000	0.3314 ± 0.0000	0.5750 ± 0.0000
	CAN	0.3300 ± 0.0000	0.3115 ± 0.0000	0.0947 ± 0.0000	0.3749 ± 0.0000	0.2403 ± 0.0000	0.8517 ± 0.0000
	LT-MSC	0.9900 ± 0.0000	0.9652 ± 0.0000	0.9750 ± 0.0000	0.9799 ± 0.0000	0.9798 ± 0.0000	0.9801 ± 0.0000
	MVCC	0.8560 ± 0.0000	0.7466 ± 0.0000	0.7137 ± 0.0000	0.7713 ± 0.0000	0.7599 ± 0.0000	0.7831 ± 0.0000
	MVGL	0.9080 ± 0.0000	0.8097 ± 0.0000	0.7742 ± 0.0000	0.8199 ± 0.0000	0.8024 ± 0.0000	0.8383 ± 0.0000
	MLRSSC	0.7080 ± 0.0000	0.7660 ± 0.0000	0.6752 ± 0.0000	0.8851 ± 0.0000	0.8300 ± 0.0000	0.6980 ± 0.0000
NGs	CSMSC	0.9840 ± 0.0000	0.9461 ± 0.0000	0.9603 ± 0.0000	0.9682 ± 0.0000	0.9681 ± 0.0000	0.9683 ± 0.0000
	GSF	0.5960 ± 0.0000	0.4409 ± 0.0000	0.3517 ± 0.0000	0.4945 ± 0.0000	0.4452 ± 0.0000	0.5562 ± 0.0000
	GMC	0.9820 ± 0.0000	0.9392 ± 0.0000	0.9554 ± 0.0000	0.9623 ± 0.0000	0.9642 ± 0.0000	0.9643 ± 0.0000
	CGD	0.9780 ± 0.0000	0.9253 ± 0.0000	0.9457 ± 0.0000	0.9565 ± 0.0000	0.9564 ± 0.0000	0.9565 ± 0.0000
	CGL	0.9320 ± 0.0000	0.8143 ± 0.0000	0.8373 ± 0.0000	0.8696 ± 0.0000	0.8688 ± 0.0000	0.8704 ± 0.0000
	CDMGC	0.7760 ± 0.0000	0.8287 ± 0.0000	0.7126 ± 0.0000	0.7798 ± 0.0000	0.6616 ± 0.0000	0.9494 ± 0.0000
	LRTGFL	**0.9940 ± 0.0000**	**0.9832 ± 0.0000**	**0.9851 ± 0.0000**	**0.9881 ± 0.0000**	**0.9879 ± 0.0000**	**0.9882 ± 0.0000**
	SSC	0.4513 ± 0.0095	0.6985 ± 0.0040	0.3052 ± 0.0051	0.3121 ± 0.0050	0.2963 ± 0.0059	0.3297 ± 0.0061
	LRR	0.6149 ± 0.0091	0.7993 ± 0.0031	0.4866 ± 0.0104	0.4917 ± 0.0102	0.4705 ± 0.0137	0.5150 ± 0.0086
	CAN	0.6437 ± 0.0000	0.8362 ± 0.0000	0.3495 ± 0.0000	0.3584 ± 0.0000	0.2458 ± 0.0000	0.6613 ± 0.0000
	RMSC	0.7849 ± 0.0102	0.9121 ± 0.0053	0.7214 ± 0.0125	0.7241 ± 0.0124	0.6864 ± 0.0117	0.7664 ± 0.0151
	LT-MSC	0.7284 ± 0.0203	0.8674 ± 0.0089	0.6333 ± 0.0231	0.6369 ± 0.0229	0.6059 ± 0.0262	0.6714 ± 0.0202
	MVCC	0.2713 ± 0.0000	0.6445 ± 0.0000	0.2108 ± 0.0000	0.2224 ± 0.0000	0.1417 ± 0.0000	0.5159 ± 0.0000
	MVGL	0.7062 ± 0.0000	0.8432 ± 0.0000	0.2795 ± 0.0000	0.2093 ± 0.0000	0.1824 ± 0.0000	0.7100 ± 0.0000
100leaves	MLRSSC	0.5256 ± 0.0000	0.4092 ± 0.0000	0.3823 ± 0.0000	0.4401 ± 0.0000	0.7723 ± 0.0000	0.4032 ± 0.0000
	CSMSC	0.6019 ± 0.0111	0.7916 ± 0.0047	0.4766 ± 0.0109	0.4817 ± 0.0108	0.4592 ± 0.0109	0.5066 ± 0.0115
	GSF	0.6762 ± 0.0000	0.8418 ± 0.0000	0.4895 ± 0.0000	0.4956 ± 0.0000	0.3929 ± 0.0000	0.6710 ± 0.0000
	GMC	0.8690 ± 0.0034	0.9436 ± 0.0001	0.7915 ± 0.0059	0.7936 ± 0.0113	0.7306 ± 0.0113	0.8687 ± 0.0037
	CGD	0.7737 ± 0.0000	0.9007 ± 0.0000	0.6790 ± 0.0000	0.6826 ± 0.0000	0.5767 ± 0.0000	0.8361 ± 0.0000
	CGL	0.9583 ± 0.0107	0.9800 ± 0.0020	0.9368 ± 0.0120	0.9375 ± 0.0119	0.9213 ± 0.0203	0.9544 ± 0.0034
	CDMGC	0.8612 ± 0.0000	0.9372 ± 0.0000	0.5377 ± 0.0000	0.5437 ± 0.0000	0.3918 ± 0.0000	0.8876 ± 0.0000
	LRTGFL	**0.9614 ± 0.0062**	**0.9913 ± 0.0015**	**0.9564 ± 0.0063**	**0.9564 ± 0.0062**	**0.9568 ± 0.0103**	**0.9821 ± 0.0040**

$$\min_{\boldsymbol{Z}^{(v)}} \sum_{v=1}^{V} \sum_{i,j=1}^{N} \|\boldsymbol{x}_i^{(v)} - \boldsymbol{x}_j^{(v)}\|_2^2 z_{ij}^{(v)m} - r \sum_{v=1}^{V} \|\boldsymbol{Z}^{(v)}\|_1 + \lambda \|\boldsymbol{\mathcal{Z}}\|_{\circledast},$$
$$s.t. \sum_{i=1}^{N} z_{ij}^{(v)} = 1, z_{ij}^{(v)} \geq 0, \tag{20}$$

b)The effectiveness of t-SVD based nuclear norm

The t-SVD based nuclear norm is deleted to verify its effectiveness in this ablate method. This ablate method is called as graph fuzzy learning (GFL). The objective function of GFL is as follows:

$$\min_{\boldsymbol{Z}^{(v)}} \sum_{v=1}^{V} \sum_{i,j=1}^{N} D_{JS}(\boldsymbol{x}_i^{(v)} || \boldsymbol{x}_j^{(v)}) z_{ij}^{(v)m} - r \sum_{v=1}^{V} \|\boldsymbol{Z}^{(v)}\|_1,$$
$$s.t. \sum_{i=1}^{N} z_{ij}^{(v)} = 1, z_{ij}^{(v)} \geq 0, \tag{21}$$

TABLE IV
p-VALUES BETWEEN THE PROPOSED LRTGFL AND OTHER CLUSTERING METHODS.

Datesets	Method	ACC	NMI	AR	F-score	Precision	Recall
ORL	SSC	3.2671e-17	1.2556e-21	4.2495e-19	4.2724e-19	3.3613e-15	7.7572e-23
	LRR	7.3335e-16	3.7292e-19	1.6529e-19	1.6203e-19	1.6429e-17	2.8437e-18
	CAN	2.6985e-13	1.8521e-16	7.2953e-16	8.2958e-16	7.1078e-15	7.8534e-15
	RMSC	5.5614e-15	1.1005e-18	4.2792e-18	4.3675e-18	1.8001e-15	3.0516e-19
	LT-MSC	1.5930e-11	2.1310e-13	3.0790e-13	2.9819e-13	2.7915e-12	6.6616e-13
	MVCC	3.6675e-15	1.2575e-18	3.1351e-16	3.6294e-16	4.1652e-15	2.2629e-15
	MVGL	7.7539e-13	1.7360e-15	1.2117e-15	1.3813e-15	9.3466e-15	3.5008e-13
	CSMSC	9.3745e-13	1.2028e-18	7.4439e-16	7.7779e-16	3.1030e-14	1.2254e-18
	GSF	2.0981e-09	2.3896e-12	2.8094e-11	2.8392e-11	1.5689e-10	2.6314e-12
	GMC	4.2222e-12	6.3787e-18	2.5358e-20	4.1339e-20	9.6991e-18	7.8728e-15
	CGD	1.2138e-11	6.0111e-15	8.9478e-14	9.4480e-14	3.1608e-13	1.8598e-13
	CGL	4.8894e-08	1.9898e-14	5.8428e-11	5.8304e-11	8.1930e-09	7.3300e-15
	CDMGC	2.7639e-12	8.0016e-15	1.5063e-15	1.7183e-15	1.0498e-14	2.8653e-12
Yale	SSC	2.9893e-08	2.3079e-13	7.0520e-10	6.6975e-10	1.0851e-08	9.6048e-12
	LRR	2.9893e-08	2.3079e-13	7.0520e-10	6.6975e-10	1.0851e-08	9.6048e-12
	CAN	5.1406e-09	3.4816e-12	1.4593e-10	1.8256e-10	4.5241e-10	3.6871e-10
	RMSC	5.4035e-09	1.4639e-12	2.7556e-11	5.8983e-11	7.0238e-11	6.0675e-12
	LT-MSC	5.2955e-05	2.5836e-09	3.1027e-07	3.1628e-07	1.6610e-06	7.3310e-08
	MVCC	5.4035e-09	1.4639e-12	2.7556e-11	5.8983e-11	7.0238e-11	6.0675e-12
	MVGL	1.1315e-07	7.1564e-11	4.4544e-09	4.6695e-09	1.5855e-08	1.6856e-09
	CSMSC	1.2045e-07	8.7700e-11	2.2479e-08	2.2412e-08	1.2577e-07	2.3153e-09
	GSF	2.6557e-07	4.0644e-10	2.5709e-08	2.5769e-08	1.2594e-08	5.0716e-09
	GMC	8.7133e-07	3.0100e-10	8.1248e-09	8.7917e-09	1.9866e-08	7.0918e-09
	CGD	1.7167e-07	8.2656e-11	9.8807e-09	9.9150e-09	5.0746e-08	1.8588e-09
	CGL	1.5981e-04	2.2475e-08	3.6195e-06	3.5038e-06	2.8556e-05	3.4673e-07
	CDMGC	1.1697e-05	3.1795e-09	2.8296e-09	3.3022e-09	5.0318e-09	1.1775e-08
UCI	SSC	1.7574e-28	2.0810e-26	4.4404e-28	4.2884e-28	3.6424e-28	5.2618e-28
	LRR	1.7706e-20	1.9393e-17	3.6683e-19	8.3655e-19	9.4607e-23	3.2662e-13
	CAN	0	1.9175e-253	0	9.2941e-138	0	4.2788e-265
	RMSC	2.3798e-23	1.4840e-25	1.6726e-23	5.0851e-23	5.0851e-23	4.4047e-24
	LT-MSC	7.8528e-14	5.7468e-14	1.8581e-13	1.8865e-13	7.1012e-14	5.2352e-13
	MVCC	0	1.6937e-233	0	0	0	8.1063e-208
	MVGL	1.0803e-137	2.3951e-194	7.3585e-139	0	0	4.9462e-137
	CSMSC	3.7928e-18	6.0701e-21	4.1289e-19	4.0293e-19	6.3929e-19	2.3633e-19
	GSF	6.9811e-142	2.1769e-171	2.8800e-146	1.1838e-145	2.2551e-146	3.5382e-273
	GMC	0	1.1688e-245	0	3.7631e-141	4.1852e-142	1.9854e-271
	CGD	0	2.8508e-136	1.5173e-139	0	2.1234e-139	4.7696e-271
	CGL	**0.0552**	5.8495e-06	0.0041	0.0039	0.0171	1.8622e-05
	CDMGC	0	1.8870e-233	0	1.6662e-140	0	1.6078e-138
Handwritten	LRR	8.0311e-28	4.1986e-26	3.6880e-26	3.9456e-26	1.2883e-26	1.4391e-25
	CAN	0	1.6303e-137	0	1.9920e-140	2.6450e-141	0
	RMSC	1.5741e-19	8.0998e-20	5.9176e-20	5.8177e-20	9.4312e-20	3.6048e-20
	LT-MSC	4.2540e-09	1.1035e-10	3.6757e-10	3.6463e-10	4.2707e-10	3.0822e-10
	MVGL	0	3.7427e-134	8.6591e-138	0	9.6943e-139	0
	CSMSC	1.3040e-21	1.3306e-24	2.7777e-22	2.7237e-22	4.9832e-22	1.3609e-22
	GMC	0	2.0314e-202	0	3.8174e-142	1.6239e-142	1.2640e-141
	CGD	2.4901e-137	0	0	2.0706e-137	0	0
	CGL	3.6806e-130	6.1891e-134	0	8.0789e-133	7.2105e-133	9.0662e-133
	CDMGC	1.3608e-137	5.5997e-132	3.1518e-135	1.6989e-137	9.5250e-139	0

c)The effectiveness of fuzzification

In this case, fuzzification factor is set as 1 to verify the effectiveness of fuzzification. This ablate method is called as low-rank tensor regularized graph learning (LRTGL). The objective function of NLRTGL is as follows:

$$\min_{\boldsymbol{Z}^{(v)}} \sum_{v=1}^{V} \sum_{i,j=1}^{N} D_{JS}(\boldsymbol{x}_i^{(v)} \| \boldsymbol{x}_j^{(v)}) z_{ij}^{(v)} - r \sum_{v=1}^{V} \|\boldsymbol{Z}^{(v)}\|_1 + \lambda \|\boldsymbol{\mathcal{Z}}\|_{\circledast},$$
$$s.t. \sum_{i=1}^{N} z_{ij}^{(v)} = 1, z_{ij}^{(v)} \geq 0, \tag{22}$$

Tables VI shows the comparison of LRTGFL, LRTGFL-ED, GFL and LRTGL on eight real datasets. It is obvious that all of the Jensen-Shannon divergence, t-SVD based nuclear norm and fuzzification improve the performance of LRTGFL. t-SVD based nuclear norm can expose the high-dimension information between views. It is most influential in ORL, UCI, Handwritten, NGs and 100leaves. In fact, it promotes to find the relationships between views thus it enhance the performance of the algorithm greatly on all of the eight datasets. Fuzzification acts on each view and makes the algorithm more softer to reduce misclassification. It performs most influentially in Yale, BBCsport and BBC4view. Jenson-Shannon divergence can obtain more nonlinear structures in data and it is better than Euclidean distance in the experiments, but it might not adjust to all views of a dataset. Thus compared with the other two components, the Jensen-Shannon divergence corresponds to a relatively small boost to the algorithm.

VI. Conclusion

In the paper, a graph fuzzy learning method for multi-view data processing is proposed. Firstly, the Jensen-Shannon divergence is adopted to represent the distance between data

TABLE V
p-VALUES BETWEEN THE PROPOSED LRTGFL AND OTHER CLUSTERING METHODS.

Datesets	Method	ACC	NMI	AR	F-score	Precision	Recall
BBCsport	SSC	1.2718e-21	6.0294e-24	1.6320e-22	5.9872e-23	8.1598e-22	7.8563e-25
	LRR	1.6348e-133	3.2686e-138	0	0	0	1.2229e-136
	CAN	0	7.9752e-141	0	4.6410e-141	0	6.5190e-132
	RMSC	0	1.4607e-142	1.0461e-137	0	6.5238e-137	2.5657e-136
	LT-MSC	3.3008e-11	8.0811e-14	1.4645e-10	5.4911e-12	9.3079e-11	1.5056e-06
	MVCC	0	0	0	1.3771e-137	7.1399e-139	3.1685e-135
	MVGL	0	4.2802e-136	0	3.5719e-134	0	9.3446e-134
	CSMSC	0	1.6254e-137	4.3131e-137	0	1.7964e-133	3.8413e-133
	GSF	2.2735e-139	1.0235e-139	3.4125e-141	0	1.9398e-141	5.6678e-133
	GMC	1.1783e-139	6.5571e-139	2.5792e-141	0	1.3988e-141	0
	CGD	1.3388e-130	1.1387e-132	1.7475e-134	0	0	0
	CGL	7.1952e-131	4.7794e-138	3.1504e-137	0	7.0615e-132	2.8028e-137
	CDMGC	0	8.7900e-140	0	0	1.3108e-141	2.4605e-128
BBC4view	SSC	0	9.8468e-192	0	1.0189e-141	2.5188e-142	9.8360e-141
	LRR	8.4714e-22	3.2366e-22	7.5618e-24	9.7140e-24	1.1250e-24	9.7358e-23
	CAN	0	5.7669e-187	1.2611e-152	2.0027e-145	0	0
	RMSC	1.1668e-139	4.57552e-218	5.0594e-145	2.8441e-141	2.6937e-141	3.0022e-141
	LT-MSC	7.9686e-15	1.9088e-18	5.6241e-19	8.7967e-19	5.6284e-20	5.1475e-18
	MVCC	3.7983e-139	9.1618e-261	0	0	4.4027e-140	2.3476e-139
	MVGL	0	5.4524e-248	0	0	5.5779e-142	0
	CSMSC	2.3135e-136	1.6597e-268	0	0	1.6711e-138	0
	GSF	0	1.3927e-192	1.9619e-152	1.0034e-145	0	0
	GMC	3.2952e-140	8.5835e-256	3.1019e-142	7.5025e-141	4.1369e-142	7.0745e-135
	CGD	2.1675e-133	9.6455e-258	8.0012e-139	1.0988e-137	6.4761e-139	1.4904e-135
	CGL	0	1.1468e-236	9.8114e-140	9.4126e-139	0	0
	CDMGC	5.5974e-145	9.5139e-176	1.4650e-149	1.9857e-142	0	0
NGs	SSC	2.0250e-141	1.3777e-140	7.8400e-146	4.7501e-142	1.0816e-142	8.1916e-141
	LRR	0	3.5600e-161	0	0	2.9285e-143	1.8860e-141
	CAN	0	3.1859e-120	1.4044e-152	0	3.9023e-209	1.1696e-267
	RMSC	1.0782e-21	1.0364e-22	8.0787e-25	1.1476e-22	3.5172e-26	1.6304e-17
	LT-MSC	0	1.0180e-217	5.9027e-127	4.3211e-126	1.3610e-245	3.7145e-179
	MVCC	3.6522e-137	1.8570e=231	0	0	1.1579e-271	7.6928e-271
	MVGL	0	2.2009e-229	0	6.1809e-138	4.7203e-270	1.7307e-137
	CSMSC	0	3.1859e-150	1.4044e-152	0	3.9023e-209	1.1696e-267
	GSF	0	2.0314e-202	0	3.8174e-142	1.6239e-142	1.2640e-141
	GMC	0	6.7920e-251	0	1.3966e-127	6.0151e-254	2.5867e-130
	CGD	0	4.7441e-222	0	0	2.1874e-131	2.0409e-131
	CGL	4.8972e-134	1.2137e-245	1.9726e-137	7.4065e-134	1.3813e-136	1.6687e-266
	CDMGC	5.9616e-139	1.4813e-197	7.9943e-140	0	1.8184e-274	3.2740e-132
100leaves	SSC	1.8235e-25	6.4991e-22	2.7213e-32	2.8121e-32	4.2752e-25	3.5220e-30
	LRR	4.5444e-23	3.1252e-23	1.0731e-23	9.4466e-24	1.5232e-23	3.5454e-22
	CAN	6.3489e-17	9.8245e-20	2.1583e-19	2.2401e-19	6.3057e-18	1.0563e-18
	RMSC	1.7568e-17	2.6853e-13	7.2612e-17	7.3128e-17	1.4973e-20	5.9326e-13
	LT-MSC	2.5560e-12	3.1642e-12	6.2044e-13	6.1154e-13	1.8264e-13	8.1080e-13
	MVCC	5.8892e-20	7.0285e-23	3.3832e-20	3.5452e-20	1.7715e-18	3.6619e-20
	MVGL	4.5585e-16	1.4091e-19	8.0825e-20	8.4890e-20	2.8497e-18	4.6536e-18
	CSMSC	7.8887e-21	1.8560e-18	5.0902e-23	5.0453e-23	4.39943e-26	8.8734e-19
	GSF	1.6767e-16	1.3664e-19	2.2864e-18	2.3317e-18	5.5053e-17	1.3943e-18
	GMC	5.4477e-16	1.3190e-21	3.1788e-22	3.1899e-22	3.0220e-19	6.8613e-23
	CGD	7.2308e-15	1.2360e-16	2.4745e-16	2.5074e-16	2.3264e-15	1.2577e-15
	CGL	**0.6058**	8.5711e-11	4.6924e-04	4.6681e-04	**0.1322**	2.7619e-12
	CDMGC	2.0398e-12	1.2778e-15	6.0853e-18	6.2815e-18	5.4129e-17	6.2753e-14

points to extract nonlinear structure. Besides, fuzzy is added to the similarity graph matrices to make the clustering softer. Furthermore, a low-rank tensor constraint is taken to use high-dimension information. The experiments on real-world datasets which compared with fourteen state-of-the-art methods demonstrated the superiority of the proposed LRTGFL. Besides, ablation study verifies the effectiveness of the components in LRTGFL.

The paper first introduces fuzzy into multi-view graph learning, and obtains a excellent performance. It points out that optimization with flexible fuzzification factor should be considered to make a softer graph clustering method in future work.

ACKNOWLEDGMENTS

The work was supported by the National Natural Science Foundation of China (Grant No.61873213), National Key Research and Development Project of China (Grant No.2018AAA0100101), the National Natural Science Foundation of China (Grant No.62003281), Natural Science Foundation of Chongqing, China (Grant No.cstc2021jcyj-msxmX1169), the Science and Technology Research Program of Chongqing Municipal Education Commission, China (Grant No.KJQN202200207).

REFERENCES

[1] J. Wang, J. Cao, B. Li, S. Lee, and R. S. Sherratt, "Bio-inspired ant colony optimization based clustering algorithm with mobile sinks for applications in consumer

TABLE VI
COMPARISON OF LRTGFL AND THREE ABLATIVE METHODS. (MEAN±STANDARD DEVIATION).

Dateset	Method	ACC	NMI	AR	F-score	Precision	Recall
ORL	LRTGFL	$\mathbf{0.9408 \pm 0.0189}$	$\mathbf{0.9849 \pm 0.0049}$	$\mathbf{0.9389 \pm 0.0185}$	$\mathbf{0.9403 \pm 0.0181}$	$\mathbf{0.9148 \pm 0.0251}$	$\mathbf{0.9675 \pm 0.0112}$
	LRTGFL-ED	0.8592 ± 0.0231	0.9392 ± 0.0083	0.8133 ± 0.0262	08178 ± 0.0255	0.7740 ± 0.0299	0.8669 ± 0.0227
	GFL	0.6473 ± 0.0274	0.8194 ± 0.0087	0.5342 ± 0.0259	0.5457 ± 0.0250	0.4998 ± 0.0333	0.6016 ± 0.0133
	LRTGL	0.7427 ± 0.0152	0.8866 ± 0.0061	0.6543 ± 0.0202	0.6630 ± 0.0196	0.5967 ± 0.0268	0.7465 ± 0.0149
Yale	LRTGFL	$\mathbf{0.8582 \pm 0.0544}$	$\mathbf{0.9068 \pm 0.0234}$	$\mathbf{0.8043 \pm 0.0567}$	$\mathbf{0.8167 \pm 0.0530}$	$\mathbf{0.7913 \pm 0.0644}$	$\mathbf{0.8445 \pm 0.0434}$
	LRTGFL-ED	0.6974 ± 0.0019	0.7200 ± 0.0029	0.5131 ± 0.0067	0.5446 ± 0.0063	0.5134 ± 0.0069	0.5799 ± 0.0054
	GFL	0.5655 ± 0.0140	$0.5817 \pm 0..0112$	0.3127 ± 0.0200	0.3618 ± 0.0171	0.3096 ± 0.0254	0.4376 ± 0.0181
	LRTGL	0.4515 ± 0.0205	0.4636 ± 0.0181	0.1796 ± 0.0196	0.2330 ± 0.0179	0.2185 ± 0.0189	0.2497 ± 0.0172
UCI	LRTGFL	$\mathbf{0.9980 \pm 0.0000}$	$\mathbf{0.9948 \pm 0.0000}$	$\mathbf{0.9956 \pm 0.0000}$	$\mathbf{0.9960 \pm 0.0000}$	$\mathbf{0.9960 \pm 0.0000}$	$\mathbf{0.9960 \pm 0.0000}$
	LRTGFL-ED	0.8550 ± 0.0000	0.9120 ± 0.0000	0.8384 ± 0.0000	0.8556 ± 0.0000	0.7996 ± 0.0000	0.9200 ± 0.0000
	GFL	0.8136 ± 0.0105	0.7963 ± 0.0067	0.7315 ± 0.0039	0.7597 ± 0.0035	0.7196 ± 0.0031	0.8045 ± 0.0048
	LRTGL	0.8875 ± 0.0000	0.9182 ± 0.0000	0.8645 ± 0.0000	0.8787 ± 0.0000	0.8328 ± 0.0000	0.9300 ± 0.0000
Handwritten	LRTGFL	$\mathbf{1.0000 \pm 0.0000}$	$\mathbf{1.0000 \pm 0.0000}$	$\mathbf{1.0000 \pm 0.0000}$	$\mathbf{1.0000 \pm 0.0000}$	$\mathbf{1.0000 \pm 0.0000}$	$\mathbf{1.0000 \pm 0.0000}$
	LRTGFL-ED	0.8625 ± 0.0000	0.9240 ± 0.0000	0.8517 ± 0.0000	0.8676 ± 0.0000	0.8088 ± 0.0000	0.9356 ± 0.0000
	GFL	0.7964 ± 0.0081	0.8046 ± 0.0029	0.7290 ± 0.0029	0.7574 ± 0.0024	0.7197 ± 0.0067	0.7994 ± 0.0046
	LRTGL	0.8635 ± 0.0000	0.9287 ± 0.0000	0.8540 ± 0.0000	0.8696 ± 0.0000	0.8095 ± 0.0000	0.9393 ± 0.0000
BBCsport	LRTGFL	$\mathbf{1.0000 \pm 0.0000}$	$\mathbf{1.0000 \pm 0.0000}$	$\mathbf{1.0000 \pm 0.0000}$	$\mathbf{1.0000 \pm 0.0000}$	$\mathbf{1.0000 \pm 0.0000}$	$\mathbf{1.0000 \pm 0.0000}$
	LRTGFL-ED	0.9982 ± 0.0000	0.9934 ± 0.0000	0.9966 ± 0.0000	0.9974 ± 0.0000	0.9983 ± 0.0000	0.9965 ± 0.0000
	GFL	0.3691 ± 0.0012	0.0352 ± 0.0001	0.0088 ± 0.0001	0.3872 ± 0.0001	0.2419 ± 0.0000	0.9694 ± 0.0034
	LRTGL	0.3362 ± 0.0064	0.0882 ± 0.0028	0.0410 ± 0.0056	0.3302 ± 0.0017	0.2612 ± 0.0033	0.4489 ± 0.0114
BBC4view	LRTGFL	$\mathbf{0.9912 \pm 0.0000}$	$\mathbf{0.9676 \pm 0.0000}$	$\mathbf{0.9798 \pm 0.0000}$	$\mathbf{0.9845 \pm 0.0000}$	$\mathbf{0.9890 \pm 0.0000}$	$\mathbf{0.9801 \pm 0.0000}$
	LRTGFL-ED	0.9752 ± 0.0000	0.9369 ± 0.0000	0.9478 ± 0.0001	0.9599 ± 0.0001	0.9732 ± 0.0000	0.9469 ± 0.0000
	GFL	0.4250 ± 0.0043	0.2350 ± 0.0081	0.0479 ± 0.0065	0.3793 ± 0.0065	0.2556 ± 0.0030	0.7353 ± 0.0028
	LRTGL	0.3314 ± 0.0000	0.0326 ± 0.0000	0.0038 ± 0.0000	0.3746 ± 0.0000	0.2331 ± 0.0000	0.9356 ± 0.0000
NGs	LRTGFL	$\mathbf{0.9940 \pm 0.0000}$	$\mathbf{0.9832 \pm 0.0000}$	$\mathbf{0.9214 \pm 0.0000}$	$\mathbf{0.9370 \pm 0.0000}$	$\mathbf{0.9363 \pm 0.0000}$	$\mathbf{0.9377 \pm 0.0000}$
	LRTGFL-ED	0.9680 ± 0.0000	0.9086 ± 0.0000	0.9214 ± 0.0000	0.9370 ± 0.0000	0.9363 ± 0.0000	0.9377 ± 0.0000
	GFL	0.2990 ± 0.0077	0.1807 ± 0.0114	0.0143 ± 0.0011	0.3264 ± 0.0000	0.2045 ± 0.0000	0.8080 ± 0.0086
	LRTGL	0.3620 ± 0.0000	0.1664 ± 0.0000	0.0664 ± 0.0000	0.3239 ± 0.0000	0.2330 ± 0.0000	0.5315 ± 0.0000
100leaves	LRTGFL	$\mathbf{0.9614 \pm 0.0062}$	$\mathbf{0.9913 \pm 0.0015}$	$\mathbf{0.9564 \pm 0.0063}$	$\mathbf{0.9564 \pm 0.0062}$	$\mathbf{0.9568 \pm 0.0103}$	$\mathbf{0.9821 \pm 0.0040}$
	LRTGFL-ED	0.9210 ± 0.0111	0.9744 ± 0.0040	0.9049 ± 0.0130	0.9058 ± 0.0128	0.8745 ± 0.0145	0.9309 ± 0.0114
	GFL	0.8859 ± 0.0117	0.9499 ± 0.0027	0.8410 ± 0.0094	0.8426 ± 0.0093	0.8048 ± 0.0153	0.8842 ± 0.0057
	LRTGL	0.9088 ± 0.0126	0.9692 ± 0.0028	0.8868 ± 0.0125	0.8879 ± 0.0124	0.8498 ± 0.0170	0.9297 ± 0.0088

home automation networks," *IEEE Transactions on Consumer Electronics*, vol. 61, no. 4, pp. 438–444, 2015.

[2] J. Wang, Y. Yin, J. Zhang, S. Lee, and R. S. Sherratt, "Mobility based energy efficient and multi-sink algorithms for consumer home networks," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 1, pp. 77–84, 2013.

[3] J. Dombi, T. Jónás, and Z. E. Tóth, "Clustering empirical failure rate curves for reliability prediction purposes in the case of consumer electronic products," *Quality and Reliability Engineering International*, vol. 32, no. 3, pp. 1071–1083, 2016.

[4] F. Nie, X. Wang, and H. Huang, "Clustering and projected clustering with adaptive neighbors," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2014, pp. 977–986.

[5] F. Nie, J. Li, X. Li *et al.*, "Self-weighted multiview clustering with multiple graphs." in *International Joint Conference on Artificial Intelligence (IJCAI)*, 2017, pp. 2564–2570.

[6] F. Nie, G. Cai, and X. Li, "Multi-view clustering and semi-supervised classification with adaptive neighbours," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, 2017.

[7] F. Nie, J. Li, X. Li *et al.*, "Parameter-free auto-weighted multiple graph learning: a framework for multiview clustering and semi-supervised classification." in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, 2016, pp. 1881–1887.

[8] J. B. Tenenbaum, V. d. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[9] H. Che and J. Wang, "A nonnegative matrix factorization algorithm based on a discrete-time projection neural network," *Neural Networks*, vol. 103, pp. 63–71, 2018.

[10] H. Che, J. Wang, and A. Cichocki, "Bicriteria sparse nonnegative matrix factorization via two-timescale duplex neurodynamic optimization," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–11, 2021, doi:10.1109/TNNLS.2021.3125457.

[11] J. Liu, C. Wang, J. Gao, and J. Han, "Multi-view clustering via joint nonnegative matrix factorization," in *Proceedings of the 2013 SIAM International Conference on Data Mining*. SIAM, 2013, pp. 252–260.

[12] C. Li, H. Che, M.-F. Leung, C. Liu, and Z. Yan, "Robust multi-view non-negative matrix factorization with adaptive graph and diversity constraints," *Information Sciences*, vol. 634, pp. 587–607, 2023.

[13] H. Wang, Y. Yang, and T. Li, "Multi-view clustering via concept factorization with local manifold regularization," in *2016 IEEE 16th International Conference on Data*

Mining (ICDM). IEEE, 2016, pp. 1245–1250.

[14] H. Zhao, Z. Ding, and Y. Fu, "Multi-view clustering via deep matrix factorization," in *Thirty-first AAAI Conference on Artificial Intelligence*, 2017.

[15] X. Cao, C. Zhang, H. Fu, S. Liu, and H. Zhang, "Diversity-induced multi-view subspace clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 586–594.

[16] S. Luo, C. Zhang, W. Zhang, and X. Cao, "Consistent and specific multi-view subspace clustering," in *Thirty-second AAAI Conference on Artificial Intelligence*, 2018.

[17] X. Wang, X. Guo, Z. Lei, C. Zhang, and S. Z. Li, "Exclusivity-consistency regularized multi-view subspace clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 923–931.

[18] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

[19] H. Che, B. Pan, M.-F. Leung, Y. Cao, and Z. Yan, "Tensor factorization with sparse and graph regularization for fake news detection on social networks," *IEEE Transactions on Computational Social Systems*, pp. 1–11, 2023, doi:10.1109/TCSS.2023.3296479.

[20] C. Zhang, H. Fu, S. Liu, G. Liu, and X. Cao, "Low-rank tensor constrained multiview subspace clustering," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1582–1590.

[21] Y. Xie, D. Tao, W. Zhang, Y. Liu, L. Zhang, and Y. Qu, "On unifying multi-view self-representations for clustering by tensor multi-rank minimization," *International Journal of Computer Vision*, vol. 126, no. 11, pp. 1157–1179, 2018.

[22] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 1, pp. 148–172, 2013.

[23] B. Pan, C. Li, and H. Che, "Nonconvex low-rank tensor approximation with graph and consistent regularizations for multi-view subspace learning," *Neural Networks*, vol. 161, pp. 638–658, 2023.

[24] K. Zhan, C. Zhang, J. Guan, and J. Wang, "Graph learning for multiview clustering," *IEEE Transactions on Cybernetics*, vol. 48, no. 10, pp. 2887–2895, 2017.

[25] H. Wang, Y. Yang, and B. Liu, "Gmc: Graph-based multi-view clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 6, pp. 1116–1129, 2019.

[26] C. Tang, X. Liu, X. Zhu, E. Zhu, Z. Luo, L. Wang, and W. Gao, "Cgd: Multi-view clustering via cross-view graph diffusion," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 04, 2020, pp. 5924–5931.

[27] S. Huang, I. W. Tsang, Z. Xu, and J. Lv, "Measuring diversity in graph learning: a unified framework for structured multi-view clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 12, pp. 5869–5883, 2021.

[28] J. Wu, X. Xie, L. Nie, Z. Lin, and H. Zha, "Unified graph and low-rank tensor learning for multi-view clustering," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 04, 2020, pp. 6388–6395.

[29] Z. Li, C. Tang, X. Liu, X. Zheng, W. Zhang, and E. Zhu, "Consensus graph learning for multi-view clustering," *IEEE Transactions on Multimedia*, vol. 24, pp. 2461–2472, 2021.

[30] Y. Mei, Z. Ren, B. Wu, T. Yang, and Y. Shao, "Multi-order similarity learning for multi-view spectral clustering," *Pattern Recognition*, vol. 137, p. 109264, 2023.

[31] R. Wang, F. Nie, Z. Wang, H. Hu, and X. Li, "Parameter-free weighted multi-view projected clustering with structured graph learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 10, pp. 2014–2025, 2019.

[32] Q. Gao, Z. Wan, Y. Liang, Q. Wang, Y. Liu, and L. Shao, "Multi-view projected clustering with graph learning," *Neural Networks*, vol. 126, pp. 335–346, 2020.

[33] T. Liu, C. K. L. Lekamalage, G.-B. Huang, and Z. Lin, "An adaptive graph learning method based on dual data representations for clustering," *Pattern Recognition*, vol. 77, pp. 126–139, 2018.

[34] Z. Ren, Q. Sun, and D. Wei, "Multiple kernel clustering with kernel k-means coupled graph tensor learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 11, 2021, pp. 9411–9418.

[35] W. Xia, S. Wang, M. Yang, Q. Gao, J. Han, and X. Gao, "Multi-view graph embedding clustering network: Joint self-supervision and block diagonal representation," *Neural Networks*, vol. 145, pp. 1–9, 2022.

[36] J. Tang, D. Li, Y. Tian, and D. Liu, "Multi-view learning based on nonparallel support vector machine," *Knowledge-Based Systems*, vol. 158, pp. 94–108, 2018.

[37] Y.-H. H. Tsai, Y. Wu, R. Salakhutdinov, and L.-P. Morency, "Self-supervised learning from a multi-view perspective," *arXiv preprint arXiv:2006.05576*, 2020.

[38] M. Yang, C. Deng, and F. Nie, "Adaptive-weighting discriminative regression for multi-view classification," *Pattern Recognition*, vol. 88, pp. 236–245, 2019.

[39] X. Yang, H. Che, M.-F. Leung, and C. Liu, "Adaptive graph nonnegative matrix factorization with the self-paced regularization," *Applied Intelligence*, pp. 1–18, 2022.

[40] S. Wang, S. Xiao, W. Zhu, and Y. Guo, "Multi-view fuzzy clustering of deep random walk and sparse low-rank embedding," *Information Sciences*, vol. 586, pp. 224–238, 2022.

[41] J. Wu, Z. Lin, and H. Zha, "Essential tensor learning for multi-view spectral clustering," *IEEE Transactions on Image Processing*, vol. 28, no. 12, pp. 5910–5922, 2019.

[42] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[43] F. S. Samaria and A. C. Harter, "Parameterisation of a stochastic model for human face identification," in

Proceedings of 1994 IEEE Workshop on Applications of Computer Vision. IEEE, 1994, pp. 138–142.

[44] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2765–2781, 2013.

[45] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 171–184, 2012.

[46] R. Xia, Y. Pan, L. Du, and J. Yin, "Robust multi-view spectral clustering via low-rank and sparse decomposition," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 28, no. 1, 2014.

[47] M. Brbić and I. Kopriva, "Multi-view low-rank sparse subspace clustering," *Pattern Recognition*, vol. 73, pp. 247–258, 2018.

[48] K. Zhan, C. Niu, C. Chen, F. Nie, C. Zhang, and Y. Yang, "Graph structure fusion for multiview clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 10, pp. 1984–1993, 2018.

Baicheng Pan received the B.S. degree in Huazhong University of Science and Technology, Wuhan, China, in 2019. He is currently pursuing the M.S. degree with the College of Electronic and Information Engineering, Southwest University, Chongqing, China.

His research interests include optimization theory and application, machine learning, multi-view clustering, non-negative matrix factorization, tensor factorization and deep matrix factorization.

Chuandong Li (Senior Member, IEEE) received the B.S. degree in applied mathematics from Sichuan University, Chengdu, China, in 1992, and the M.S. degree in operational research and control theory, and the Ph.D. degree in computer software and theory from Chongqing University, Chongqing, China, in 2001 and 2005, respectively.

From November 2006 to November 2008, he served as a Research Fellow with the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong. From October 2011 to October 2012, he served as a Visiting Professor with Texas A & M University at Qatar, Doha, Qatar. He is currently a Professor with Southwest University, Chongqing. He has published more than 400 articles in international journals. His current research interest covers computational intelligence, impulsive systems, and game theoretical control.

Hangjun Che received the B.S. degree in electronic and information engineering from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2013, the M.S. degree in signal and information processing from Southwest University, Chongqing, in 2016, and the Ph.D. degree in computer science from the City University of Hong Kong, Hong Kong, in 2019.

He is currently an Associate Professor with the College of Electronic and Information Engineering, Southwest University. His current research interests include neurodynamic optimization, multi-view learning, and non-negative matrix factorization.

Man-Fai Leung received the Ph.D. degree in computer science from the City University of Hong Kong, Hong Kong, in 2019. He is currently with the School of Computing and Information Science, Faculty of Science and Engineering, Anglia Ruskin University, Cambridge, U.K.

His current research interests include portfolio optimization, multi-objective optimization, and computational intelligence.

Keping Yu received the M.E. and Ph.D. degrees from the Graduate School of Global Information and Telecommunication Studies, Waseda University, Tokyo, Japan, in 2012 and 2016, respectively. He was a Research Associate, Junior Researcher, Researcher with the Global Information and Telecommunication Institute, Waseda University, from 2015 to 2019, 2019 to 2020, 2020 to 2022, respectively. He is currently an associate professor at Hosei University and a visiting scientist at the RIKEN Center for Advanced Intelligence Project, Japan.

Dr. Yu has hosted and participated in more than ten projects, is involved in many standardization activities organized by ITU-T and ICNRG of IRTF, and has contributed to ITU-T Standards Y.3071 and Supplement 35. He received the IEEE Outstanding Leadership Award from IEEE BigDataSE 2021, the Best Paper Award from IEEE Consumer Electronics Magazine Award 2022 (1st Place Winner), IEEE ICFTIC 2021, ITU Kaleidoscope 2020, the Student Presentation Award from JSST 2014. He has authored more than 200 peer-review research papers and books, including over 70 IEEE/ACM Transactions papers. He is an Associate Editor of IEEE Open Journal of Vehicular Technology, Journal of Intelligent Manufacturing, Journal of Circuits, Systems and Computers, and IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences. He has been a Guest Editor for more than 20 journals such as IEEE Transactions on Computational Social Systems, IEEE Journal of Biomedical and Health Informatics, and Renewable & Sustainable Energy Reviews. He served as general co-chair and publicity co-chair of the IEEE VTC2020-Spring 1st EBTSRA workshop, general co-chair of IEEE ICCC2020 2nd EBTSRA workshop, general co-chair of IEEE TrustCom2021 3nd EBTSRA workshop, session chair of IEEE ICCC2020, ITU Kaleidoscope 2016. His research interests include smart grids, information-centric networking, the Internet of Things, artificial intelligence, blockchain, and information security.