Tensorial Multi-view Clustering with Deep Anchor Graph Projection

Wei Feng[1] , **Dongyuan Wei**[2] , **Qianqian Wang**[3*] and **Bo Dong**[4]
[1]College of Information Engineering, Northwest A&F University, Yangling, China
[2]School of Computer Science and Technology, Xi'an Jiaotong University, Xi'an, China
[3]School of Telecommunications Engineering, Xidian University, Xi'an, China
[4]School of Continuing Education, Xi'an Jiaotong University, Xi'an, China
wei.feng@nwafu.edu.cn, weidongyuan@stu.xjtu.edu.cn, qqwang@xidian.edu.cn, dong.bo@mail.xjtu.edu.cn

Abstract

Multi-view clustering (MVC) has emerged as an important unsupervised multi-view learning method that leverages consistent and complementary information to enhance clustering performance. Recently, tensorized MVC, which processes multi-view data as a tensor to capture their cross-view information, has received considerable attention. However, existing tensorized MVC methods generally overlook deep structures within each view and rely on post-processing to derive clustering results, leading to potential information loss and degraded performance. To address these issues, we develop **T**ensorial **M**ulti-**v**iew **C**lustering with **D**eep **A**nchor **G**raph **P**rojection (TMVC-DAGP), which performs deep projection on the anchor graph, thus improving model scalability. Besides, we utilize a sparsity regularization to eliminate the redundancy and enforce the projected anchor graph to retain a clear clustering structure. Furthermore, TMVC-DAGP leverages weighted Tensor Schatten p-norm to exploit the consistent and complementary information. Extensive experiments on multiple datasets demonstrate TMVC-DAGP's effectiveness and superiority.

1 Introduction

Multi-view clustering (MVC) improves clustering performance by exploiting the consistent and complementary information from multiple perspectives to uncover the intrinsic data clustering structure [Fu *et al.*, 2020], which receives significant attention in various applications like bioinformatics [Yang *et al.*, 2022; Gao *et al.*, 2020b]. MVC methods can be roughly categorized into subspace learning methods, graph-based methods, co-training methods, and multi-kernel learning methods. Among them, subspace learning techniques project multi-view data into a common latent subspace for clustering [Zheng *et al.*, 2023; Liu *et al.*, 2021]; graph-based methods construct similarity graphs to represent sample relationships to capture the geometric structure and then apply spectral clustering on the graph [Wei *et al.*, 2017]; co-training approaches iteratively train models on each view and exchange information to enhance clustering results, making them particularly effective when views provide complementary information [Jiang *et al.*, 2013]; multi-kernel learning methods integrate diverse views by constructing a unified kernel space, enabling flexible representation learning and improving clustering accuracy [Tzortzis and Likas, 2012].

However, traditional approaches face several challenges, including inefficiency in handling high-dimensional or large-scale data [Li *et al.*, 2020b]. Moreover, most methods achieve view fusion by minimizing the differences between individual view embeddings and a shared global embedding, which often fails to fully leverage the complementary information across different views [Gao *et al.*, 2020b]. To address efficiency issues, several works introduce the anchor technique [Li *et al.*, 2015; Li *et al.*, 2024c; Qin *et al.*, 2024], which selects a subset of representative samples as anchors and constructs an $n \times m$ anchor graph for clustering instead of the raw data, where m is anchor number [Wang *et al.*, 2021b]. In this way, these methods significantly improve model efficiency. Some methods further represent multi-view data as high-order tensors to capture complementary information across views, which well preserves spatial structures and inter-view dependencies [Guo *et al.*, 2022]. The above methods generally rely on post-processing to obtain the final clustering labels, which increases computational complexity and may limit clustering performance. Some efforts have focused on directly projecting anchor graphs into the label space via bipartite graphs to avoid post-processing [Lei *et al.*, 2024; Zhao *et al.*, 2024]. Among them, AGFS-OMVC [Zhao *et al.*, 2024] combines feature selection with anchor graph construction by using sparse constraints to remove noisy anchors and projecting the anchors directly into the label space, and obtains clustering results without post-processing. However, this method adopts single-layer projection, which is not feasible for extracting the hierarchical structures within multi-view data.

To mitigate these problems, we propose a novel tensorized MVC method, *i.e.*, **T**ensorial **M**ulti-**v**iew **C**lustering with **D**eep **A**nchor **G**raph **P**rojection (TMVC-DAGP). The proposed method guarantees efficiency by utilizing anchor graphs and addresses the deep feature extraction problem by introducing a deep multi-layer projection strategy. Specif-

*Corresponding author

ically, inspired by the MVC method based on deep non-negative matrix factorization (NMF) [Zhao *et al.*, 2017], we design a multi-layer projection structure that is capable of extracting the deep and hierarchical latent features. Besides, it incorporates a sparsity constraint on projection matrices to enhance robustness by focusing on significant features. Additionally, tensor-based fusion with Schatten p-norm regularization is employed to capture inter-view complementary information, ensuring effective use of correlations between views. The main contributions of this work are as follows:

- We develop a novel deep multi-layer anchor graph projection method that directly maps anchor graphs into the label space via deep projection, effectively extracting deep information from multi-view data.
- This method employs a $\ell_{2,p}$-norm based sparse regularization on the projection matrices, which eliminates the negative impact of noises and redundancy and thus does not rely on post-processing for accurate clustering.
- We conduct extensive experimental validation on multiple multi-view datasets, demonstrating the superiority of TMVC-DAGP over existing approaches in clustering performance and computational efficiency.

2 Related Work

2.1 Tensor-Based Multi-view Clustering

Most of the existing MVC methods perform view fusion by minimizing the difference between view embeddings and global embeddings, and this fusion strategy does not fully utilize the view complementary information of different views of multi-view data [Zhao *et al.*, 2024]. To overcome this limitation, tensor-based MVC methods have gained increasing attention due to their ability to capture structural dependencies and complementary information across views [Gao *et al.*, 2020b]. Tensor nuclear norm minimization, particularly when combined with tensor singular value decomposition (t-SVD), has emerged as a powerful tool to model higher-order correlations and preserve the spatial structure of multi-view data [Xia *et al.*, 2021]. For instance, low-rank tensor-based proximity learning jointly optimizes multi-view affinity matrices and consensus graphs within a unified framework, effectively capturing deeper cross-view relationships [Chen *et al.*, 2022]. EMVC-NTLC applies low-rank constraints to anchor maps of different views using tensor sp norm to extract complementary information from multi-view data [Li *et al.*, 2023]. Orth-NTF extends the matrix factorization to the tensor factorization and proposes a nonnegative tensor factorization method for MVC, which directly constructs multi-view data into tensors for decomposition and effectively improves clustering results [Li *et al.*, 2024b].

2.2 Multi-view Projection for Clustering

Feature projection aims to map high-dimensional data into lower-dimensional spaces, effectively reducing noise and redundancy while preserving essential structure [Li *et al.*, 2024a]. Traditional projection-based clustering methods primarily focus on dimensionality reduction and feature selection to enhance clustering performance. Gao *et al.* [2020a] proposed a framework that combines dimensionality reduction with manifold learning, ensuring an effective balance between data representation and computational efficiency. Similarly, Wang *et al.* [2021a] introduced sparse regularization based on $\ell_{2,1}$-norm to improve the constructed graph, providing enhanced noise resistance and clustering accuracy. To further improve clustering quality and address the limitations of traditional approaches, recent studies have explored tensor-based projection techniques. Li *et al.* [2024a] proposed a label learning method based on tensor projection (LLMTP), which projects the anchor graph directly into the label space using an orthogonal projection matrix, thus enabling one-step clustering without the need for post-processing. By extending the projection from matrices to tensors, the approach effectively captures the spatial structure and complementary information across views. Additionally, the incorporation of the tensor Schatten p-norm regularization helps ensure consistency across views. Despite these advancements, existing projection-based methods often adopts single-layer projection and thus neglect the deep latent information within each individual view, leaving room for further improvements in representation learning and clustering robustness.

3 Methodology

3.1 Motivations and Problem Statement

MVC aims to obtain an indicator matrix that uncovers the underlying cluster structure by leveraging multiple data representations. However, current methods generally confronts the following challenges: (1) **insufficient deep information extraction**, as shallow structures fail to capture complex hierarchical patterns; (2) **reliance on post-processing**, which increases complexity and introduces potential information loss; (3) **ineffective exploitation of inter-view complementarity**, limiting the integration of diverse information; and (4) **sensitivity to noise**, which affects clustering robustness. Therefore, a more effective framework is needed to extract deep features, minimize post-processing, and enhance multi-view information integration while ensuring robustness.

Problem Formulation: Given multi-view dataset $\mathbf{X} = \{\mathbf{X}^v\}(1 \leq v \leq V)$, where $\mathbf{X}^v \in \mathbb{R}^{n \times d_v}$ denotes the data matrix for the v-th view, n is sample number, and d_v is the dimension of the v-th view data. Anchor-based methods reduce computational complexity by constructing an anchor matrix $\mathbf{S}^v \in \mathbb{R}^{n \times m}$ and employing $\mathbf{S}^v$ for clustering instead of $\mathbf{X}^v$, where m is the anchor number. Our objective is to derive an indicator matrix $\mathbf{H}^v \in \mathbb{R}^{n \times k}$, where k is cluster number,

3.2 Objective

To fully exploit the deep and hidden information within each view, we design a multi-layer projection structure. This structure performs layer-by-layer projection on the anchor graph, capturing complex patterns and relationships in the data. Each layer in this deep projection structure extracts features at varying layers, enhancing the overall data representation and directly projecting the data into the labeling space, resulting in more accurate clustering. The objective function

for this deep projection structure is:

$$\begin{aligned}&\min_{\mathbf{G}_i^v,\mathbf{H}_i^v}\sum_{v=1}^{V}\frac{1}{\alpha_v}\|\mathbf{S}^v\mathbf{G}_1^v\mathbf{G}_2^v\cdots\mathbf{G}_l^v-\mathbf{H}_l^v\|_F^2\\&\text{s.t.}\quad\sum_{v=1}^{V}\alpha^v=1,\quad\alpha^v\geq 0,\quad\mathbf{H}_i^v\geq 0,\quad\mathbf{H}_l^v(\mathbf{H}_l^v)^\top=I\end{aligned}\tag{1}$$

where $\mathbf{S}^v \in \mathbb{R}^{n\times m}$ is the anchor graph for the v-th view constructed via the method illustrated in [Li *et al.*, 2024b]. $\mathbf{G}_i^v \in \mathbb{R}^{d_{i-1}\times d_i}$ is the projection matrix for the i-th layer of the v-th view, where d_i is the dimension of the i-th layer, l is the set number of layers, and $d_l = k$, $d_0 = m$. $\mathbf{H}_l^v \in \mathbb{R}^{n\times k}$ represents the clustering indicator matrix for the l-th layer, with k being the number of clusters. α_v is the weight assigned to the v-th view.

Tensorial Fusion: Eq. (1) only provides the view-specific clustering results, and the final result can be obtained with a regularization term $\sum\|\mathbf{H}-\mathbf{H}_l^v\|_F^2$, as adopted in [Zhao *et al.*, 2024]. However, this regularization is too strong and may hinder exploiting complete information of different views [Li *et al.*, 2023]. Following the tensored low-rank MVC methods, we introduce a weighted tensor Schatten p-norm regularization as formulated below:

$$\begin{aligned}&\min_{\mathbf{G}^v,\mathbf{H}_i^v}\sum_{v=1}^{V}\frac{1}{\alpha_v}\big(\|\mathbf{S}^v\mathbf{G}_1^v\mathbf{G}_2^v\cdots\mathbf{G}_l^v-\mathbf{H}_l^v\|_F^2+\beta\|\boldsymbol{\mathcal{H}}\|_{\boldsymbol{\omega},S_p}^p\\&\text{s.t.}\quad\sum_{v=1}^{V}\alpha^v=1,\quad\alpha^v\geq 0,\quad\mathbf{H}_i^v\geq 0,\quad\mathbf{H}_l^v(\mathbf{H}_l^v)^\top=I\end{aligned}\tag{2}$$

where $\boldsymbol{\mathcal{H}}$, constructed based on the approach specified in the [Li *et al.*, 2024b], is the tensor formed by all the indicator matrices $\mathbf{H}_l^v$, β is a parameter balancing the tensor norm constraint, and $\|\cdot\|_{\boldsymbol{\omega},S_p}$ represents the weighted tensor Schatten S_p-norm. This formulation allows us to capture the complementary information among views, further enhancing clustering performance.

Sparse Anchor Projection: To ensure robust feature selection and enhance resistance to noise, we incorporate a sparsity constraint using the $\ell_{2,p}$-norm on the projection matrix. This constraint filters out noise and redundancy, enhancing clustering accuracy and eliminating the need of post-processing for accurate clustering. The objective function with the sparsity constraint is expressed as:

$$\begin{aligned}&\min_{\mathbf{G}_i^v,\mathbf{H}_i^v}\sum_{v=1}^{V}\frac{1}{\alpha_v}\big(\|\mathbf{S}^v\mathbf{G}_1^v\mathbf{G}_2^v\cdots\mathbf{G}_l^v-\mathbf{H}_l^v\|_F^2\\&\quad+\lambda\sum_{i=1}^{l}\|\mathbf{G}_i^v\|_{2,p}\big)+\beta\|\boldsymbol{\mathcal{H}}\|_{\boldsymbol{\omega},S_p}^p\\&\text{s.t.}\quad\sum_{v=1}^{V}\alpha^v=1,\quad\alpha^v\geq 0,\quad\mathbf{H}_i^v\geq 0,\quad\mathbf{H}_l^v(\mathbf{H}_l^v)^\top=I\end{aligned}\tag{3}$$

where λ is used to adjust the impact of the sparsity constraint.

3.3 Optimization

We use the Augmented Lagrange Multiplier (ALM) to optimize the objective Eq. (2). Specifically, we introduce three auxiliary variables $\mathbf{Q}_i^v = \mathbf{G}_i^v$, $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{H}}$, and $\mathbf{M}^v = \mathbf{H}_l^v$ where $\mathbf{M}^v \geq 0$. let $\mathbf{B}_l^v = \mathbf{G}_1^v\mathbf{G}_2^v, ..., \mathbf{G}_l^v$, the objective function can be optimized by solving the following problem:

$$\begin{aligned}&\min_{\mathbf{G}_i^v,\mathbf{H}_i^v,\mathbf{Q}_i^v,\boldsymbol{\mathcal{J}},\mathbf{M}^v}\sum_{v=1}^{V}\Big(\frac{1}{\alpha_v}\big(\|\mathbf{S}^v\mathbf{B}_l^v-\mathbf{H}_l^v\|_F^2\\&\quad+\lambda\sum_{i=1}^{l}\|\mathbf{Q}_i^v\|_{2,p}\big)+\sum_{i=1}^{l}\frac{\rho_1}{2}\left\|\mathbf{G}_i^v-\mathbf{Q}_i^v+\frac{\mathbf{W}_i^v}{\rho_1}\right\|_F^2\\&\quad+\frac{\rho_2}{2}\left\|\mathbf{H}_l^v-\mathbf{M}^v+\frac{\mathbf{Y}^v}{\rho_2}\right\|_F^2\Big)+\beta\|\boldsymbol{\mathcal{J}}\|_{\boldsymbol{\omega},S_p}^p\\&\quad+\frac{\rho_3}{2}\left\|\boldsymbol{\mathcal{H}}-\boldsymbol{\mathcal{J}}+\frac{\boldsymbol{\mathcal{Z}}}{\rho_3}\right\|_F^2\\&\text{s.t.}\quad\sum_{v=1}^{V}\alpha^v=1,\quad\alpha^v\geq 0,\quad\mathbf{H}_l^v(\mathbf{H}_l^v)^\top=I,\end{aligned}\tag{4}$$

where $\boldsymbol{\mathcal{Z}} \in \mathbb{R}^{n\times k\times v}$, $\mathbf{Y}^v \in \mathbb{R}^{n\times k}$, and $\mathbf{W}_i^v \in \mathbb{R}^{d_{i-1}\times d_i}$ are Lagrange multipliers; ρ_i are the penalty parameters. We adopt alterative direction minimization method to optimize Eq.(4).

Optimizing $\mathbf{Q}_i^v$: Fixing other variables, Eq.(4) becomes:

$$\min_{\mathbf{Q}_i^v}\frac{\lambda}{\rho_1\alpha_v}\|\mathbf{Q}_i^v\|_{2,p}+\frac{1}{2}\|\mathbf{Q}_i^v-\mathbf{U}_i^v\|_F^2\tag{5}$$

where $\mathbf{U}_i^v = \mathbf{G}_i^v + \frac{\mathbf{W}_i^v}{\rho_1}$. According to [Zhao *et al.*, 2024], it can be solved with the Generalized Soft-Thresholding (GST) algorithm on $\mathbf{U}_i^v$:

$$\mathbf{Q}_{i,j:}^v=\sigma\cdot\frac{\mathbf{U}_{i,j:}^v}{\|\mathbf{U}_{i,j:}^v\|_2}\tag{6}$$

where $\mathbf{Q}_{i,j:}^v$ and $\mathbf{U}_{i,j:}^v$ are the j-th row vectors of $\mathbf{Q}_i^v$ and $\mathbf{U}_i^v$, respectively. σ is obtained by solving:

$$\sigma=\arg\min_{x\geq 0}\left\{\frac{1}{2}(x-\|\mathbf{U}_{i,j:}^v\|)^2+\frac{\lambda}{\rho_1\alpha_v}x^p\right\}\tag{7}$$

Optimizing $\mathbf{G}_i^v$: Fixing other variables, Eq.(4) becomes:

$$\min_{\mathbf{G}_i^v}\frac{1}{\alpha_v}\left\|\mathbf{S}^v\mathbf{B}_{i-1}^v\mathbf{G}_i^v-\mathbf{H}_i^v\right\|_F^2+\frac{\rho_1}{2}\left\|\mathbf{G}_i^v-\mathbf{Q}_i^v+\frac{\mathbf{W}_i^v}{\rho_1}\right\|_F^2\tag{8}$$

Taking the derivative of Eq.(8) with respect to $\mathbf{G}_i^v$ and setting it to zero, we have:

$$\begin{aligned}&\left(\frac{2}{\alpha_v}{\mathbf{B}_{i-1}^v}^\top{\mathbf{S}^v}^\top\mathbf{S}^v\mathbf{B}_{i-1}^v+2\rho_1 I\right)\mathbf{G}_i^v\\&=\frac{2}{\alpha_v}{\mathbf{B}_{i-1}^v}^\top{\mathbf{S}^v}^\top\mathbf{H}_i^v+2\rho_1\left(\mathbf{Q}_i^v-\frac{\mathbf{W}_i^v}{\rho_1}\right)\end{aligned}\tag{9}$$

Thus, the optimal solution for $\mathbf{G}_i^v$ is:

$$\begin{aligned}\mathbf{G}_i^v=&\left(\frac{2}{\alpha_v}\mathbf{B}_{i-1}^{vT}\mathbf{S}^v\top\mathbf{S}^v\mathbf{B}_{i-1}^v+\rho_1 I\right)^{-1}\\&\left(\frac{2}{\alpha_v}{\mathbf{B}_{i-1}^v}^\top{\mathbf{S}^v}^\top\mathbf{H}_i^v+\rho_1\left(\mathbf{Q}_i^v-\frac{\mathbf{W}_i^v}{\rho_1}\right)\right)\end{aligned}\tag{10}$$

Optimizing $\mathbf{H}_i^v (i < l)$: Fixing other variables, the optimization problem becomes:

$$\min_{\mathbf{H}_i^v} \|\mathbf{S}^v\mathbf{B}_i^v - \mathbf{H}_i^v\|_F^2 \tag{11}$$

To find the optimal $\mathbf{H}_i^v$, we directly set the optimization problem to zero. After that, we get:

$$\mathbf{H}_i^v = \mathbf{S}^v\mathbf{B}_i^v \tag{12}$$

Optimizing $\mathbf{H}_l^v$: Fixing other variables, the optimization problem becomes:

$$\begin{aligned}\min_{\mathbf{H}_l^v} \sum_{v=1}^{V}\Big(\frac{1}{\alpha_v}\|\mathbf{S}^v\mathbf{B}_l^v - \mathbf{H}_l^v\|_F^2 + \frac{\rho_2}{2}\left\|\mathbf{H}_l^v - \mathbf{M}^v + \frac{\mathbf{Y}^v}{\rho_2}\right\|_F^2\Big) \\ + \frac{\rho_3}{2}\left\|\boldsymbol{\mathcal{H}} - \boldsymbol{\mathcal{J}} + \frac{\boldsymbol{\mathcal{Z}}}{\rho_3}\right\|_F^2 \\ \text{s.t.} \quad \sum_{v=1}^{V}\alpha^v = 1, \quad \alpha^v \geq 0, \quad \mathbf{H}_l^v(\mathbf{H}_l^v)^\top = I\end{aligned} \tag{13}$$

This leads to the equivalent trace maximization problem:

$$\max_{\mathbf{H}_l^v} \sum_{v=1}^{V} \text{tr}\left(\mathbf{H}_l^{v\top}\mathbf{D}^v\right) \tag{14}$$

where the auxiliary variable $\mathbf{D}^v$ is defined as $\mathbf{D}^v = \frac{2}{\alpha^v}\mathbf{S}^v\mathbf{B}_l^v + \rho_2\left(\mathbf{M}^v - \frac{\mathbf{Y}^v}{\rho_2}\right) + \rho_3\left(\boldsymbol{\mathcal{J}}^v - \frac{\boldsymbol{\mathcal{Z}}^v}{\rho_3}\right)$

To solve this problem, perform a Singular Value Decomposition (SVD) on the matrix $\mathbf{D}^v$: $\mathbf{D}^v = \mathbf{U}\Sigma\mathbf{V}^\top$. The optimal solution for $\mathbf{H}_l^{v*}$ is then given by:

$$\mathbf{H}_l^{v*} = \mathbf{U}\mathbf{V}^\top \tag{15}$$

This iterative procedure updates the matrix $\mathbf{H}_l^{v*}$ for each view v, ensuring the optimization of clustering results across all views.

Optimizing $\boldsymbol{\mathcal{J}}$: Fixing other variables, the optimization problem becomes:

$$\arg\min_{\boldsymbol{\mathcal{J}}} \beta\|\boldsymbol{\mathcal{J}}\|_{\boldsymbol{\omega},S_p}^p + \frac{\rho_3}{2}\left\|\boldsymbol{\mathcal{H}} - \boldsymbol{\mathcal{J}} + \frac{\boldsymbol{\mathcal{Z}}}{\rho_3}\right\|_F^2 \tag{16}$$

According to [Zhao *et al.*, 2024], the optimal solution for Eq.(16) is

$$\boldsymbol{\mathcal{J}}^* = \Gamma_{\frac{\beta}{\rho_3}}\left(\boldsymbol{\mathcal{H}} + \frac{\boldsymbol{\mathcal{Z}}}{\rho_3}\right) \tag{17}$$

The term $\Gamma_{\frac{\beta}{\rho_3}}$ is a generalized shrinkage operator that enforces the Schatten S_p-norm regularization on the tensor $\boldsymbol{\mathcal{H}} + \frac{\boldsymbol{\mathcal{Z}}}{\rho_3}$. It controls the rank of $\boldsymbol{\mathcal{J}}$, enhancing solution robustness. The detailed optimization process is provided in the appendix for further reference.

Optimizing $\mathbf{M}^v$: The optimization problem for $\mathbf{M}^v$ can be formulated as follows:

$$\min_{\mathbf{M}^v} \sum_{v=1}^{V}\frac{\rho_2}{2}\left\|\mathbf{H}_l^v - \mathbf{M}^v + \frac{\mathbf{Y}^v}{\rho_2}\right\|_F^2 \quad \text{s.t.} \quad \mathbf{M}^v \geq 0 \tag{18}$$

The optimal solution to this problem is given by:

$$\mathbf{M}^v = \max\left(\mathbf{H}_l^v + \frac{\mathbf{Y}^v}{\rho_2}, 0\right) \tag{19}$$

Optimization of α_v: Fixing other variables, the optimization problem becomes:

$$\arg\min_{\alpha_v} \sum_{v=1}^{V}\frac{\tau_v}{\alpha_v} \quad \text{s.t.} \quad \sum_{v=1}^{V}\alpha_v = 1, \alpha_v \geq 0 \tag{20}$$

The solution is:

$$\alpha_v = \frac{\sqrt{\tau_v}}{\sum_{v=1}^{V}\sqrt{\tau_v}} \tag{21}$$

where $\tau_v = \|\mathbf{S}^v\mathbf{B}_l^v - \mathbf{H}_l^v\|_F^2 + \lambda\sum_{i=1}^{l}\|\mathbf{G}_i^v\|_{2,p}$

Update Other Variables: Update the Lagrange multipliers by

$$\mathbf{Y}_v = \mathbf{Y}_v + \rho_2(\mathbf{H}_l^v - \mathbf{M}^v)$$

$$\mathbf{W}_i^v = \mathbf{W}_i^v + \rho_2(\mathbf{G}_i^v - \mathbf{Q}_i^v)$$

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{Z}} + \rho_3(\boldsymbol{\mathcal{H}} - \boldsymbol{\mathcal{J}})$$

The penalty parameters $\rho_i = \min(\phi_\rho\rho_i, \rho_{\max})$, where ϕ_ρ and $\rho_{\max}$ are constants.

Dataset	Views	Dimension	Samples	Clusters
BBCSport	2	3283/3183	544	5
Yale	2	1024/4096	165	11
Sonar	3	20/20/20	208	2
Vehicle Sensor	4	5/5/7/5	1594	2
NGs	3	2000/2000/2000	500	5
WebKB	2	1840/3000	1051	2
MSRC	5	24/576/512/256/254	210	7
RGB-D	2	2048/300	1449	13

Table 1: Dataset Specifications

3.4 Complexity Analysis

Our method primarily consists of two main parts: the initialization of the anchor graph $\mathbf{S}^v$ and the iterative parameter updates. The computational complexity of initializing the anchor graph $\mathbf{S}^v$ is consistent with previous works, which is $\mathcal{O}(vnmd + vnm\log(m))$, where v, m, and n represent the number of views, anchors, and samples, respectively. Here, $d = \sum_{v=1}^{v} d_v$ is the total dimensionality across all views And assume that the number of iterations of the training process is t. The computational complexity of the update process for $\mathbf{Q}_i^v$, $\mathbf{H}_i^v$, $\mathbf{H}_l^v$,$\mathbf{G}_i^v$, $\boldsymbol{\mathcal{J}}$,$\mathbf{M}^v$ are: $\mathcal{O}(vnmklt)$, $\mathcal{O}(vnmklt)$, $\mathcal{O}((nvm^2 + vnmk)t)$, $\mathcal{O}((nvm^2 + vm^3)lt)$, $\mathcal{O}((vnklog(vm) + v^2nk)t)$, $\mathcal{O}(vmkt)$. Here, l, c, and v are small constants, with n usually the largest. Despite optimizing several variables, the limited number of layers (2-3) keeps the overall computational complexity, $\mathcal{O}(vnmd + vnm^2lt)$, reasonable in practice.

Datasets	**Yale**			**BBCSport**			**Sonar**			**Vehicle Sensor**		
Metrics	**ACC**	**NMI**	**PUR**	**ACC**	**NMI**	**PUR**	**ACC**	**NMI**	**PUR**	**ACC**	**NMI**	**PUR**
FastMICE	65.46	66.06	47.04	41.91	46.00	7.90	57.21	2.69	57.21	95.62	94.63	66.00
MvLRSSC	58.79	39.20	66.09	76.63	72.36	76.63	58.65	2.27	58.65	94.96	69.91	94.96
RMSL	78.78	78.23	79.39	76.63	72.36	76.63	64.90	7.04	64.90	78.12	61.84	78.12
GMC	54.55	62.44	54.55	80.70	76.00	79.43	63.94	5.49	63.94	78.02	61.50	78.12
FPMVS-CAG	50.31	59.32	51.52	42.10	15.09	51.84	50.48	0.01	53.37	69.68	69.68	95.81
MVC-DMF-PA	15.75	16.10	20.00	73.34	52.68	76.28	15.75	16.10	20.00	50.37	62.76	50.37
AGFS-OMVC	97.57	96.69	97.57	97.06	92.89	97.06	98.07	88.04	98.07	98.87	91.83	98.87
Orth-NTF	78.18	81.90	80.00	89.15	79.49	89.52	97.11	83.80	97.11	98.05	86.23	98.05
MVC-DNTF	84.24	86.39	82.42	98.05	87.85	94.85	97.11	83.80	97.11	99.62	96.79	99.62
OURS	**98.18**	**98.47**	**98.78**	**99.08**	**97.07**	**99.08**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**

Table 2: Clustering performance comparison in terms of ACC, NMI, and PUR on Yale, BBCSport, Sonar, and Vehicle Sensor datasets.

Datasets	**NGs**			**WebKB**			**MSRC**			**RGB-D**		
Metrics	**ACC**	**NMI**	**PUR**	**ACC**	**NMI**	**PUR**	**ACC**	**NMI**	**PUR**	**ACC**	**NMI**	**PUR**
FastMICE	38.40	48.00	26.63	95.62	94.63	0.66	86.67	86.67	77.73	41.81	32.61	49.53
MvLRSSC	90.26	88.82	91.72	92.58	58.19	92.58	78.57	68.55	78.57	39.00	32.40	50.59
RMSL	9.60	86.11	94.60	60.42	1.93	78.12	27.62	8.18	31.90	12.63	2.85	26.98
GMC	97.80	92.93	97.80	84.02	25.78	84.02	24.29	6.91	26.19	40.23	33.06	46.51
FPMVS-CAG	73.80	59.23	73.80	94.96	69.91	94.96	42.86	37.68	42.86	34.50	38.73	45.47
MVC-DMF-PA	86.80	80.27	86.80	89.43	50.89	89.43	91.43	85.36	91.43	16.83	**72.25**	33.12
AGFS-OMVC	97.40	93.56	97.40	98.57	86.39	98.57	99.00	96.02	98.09	68.32	69.63	68.99
Orth-NTF	95.40	89.73	95.40	96.57	73.25	96.57	99.00	**97.80**	99.00	59.07	65.78	75.56
MVC-DNTF	97.60	93.73	97.60	95.81	71.55	95.81	**99.04**	96.02	**99.04**	63.21	71.28	**82.95**
OURS	**99.00**	**96.80**	**99.00**	**99.81**	**97.47**	**99.81**	**99.04**	**97.80**	**99.04**	**72.25**	71.17	78.05

Table 3: Clustering performance comparison in terms of ACC(%), NMI(%), and PUR(%) on NGs, WebKB, MSRC, and RGB-D datasets.

4 Experiments

4.1 Dataset

We evaluate the performance of the proposed method on eight widely adapted multi-view learning benchmark datasets, which are **Yale** [Yale University, 2001], **BBCSport** [Greene and Cunningham, 2006], **Sonar** [Sejnowski and Gorman,], **Vehicle Sensor** [Duarte and Hu, 2004], **NGs** [Hussain *et al.*, 2010],**WebKB** [Blum and Mitchell, 1998],**MSRC** [Winn and Jojic, 2005] and **SentencesNYU v2 (RGB-D)** [Silberman *et al.*, 2012]. Detailed information on dataset specifications is provided in Table 1.

4.2 Comparison Methods

In this paper, we compare our proposed method with multiple state-of-the-art multi-view clustering algorithms. A detailed description of each comparison method is provided in the Appendix. The comparison methods include GMC [Wang *et al.*, 2019] , UDBGL [Fang *et al.*, 2023], DiMSC [Cao *et al.*, 2015], MvLRSSC [Brbić and Kopriva, 2018], RMSL [Li *et al.*, 2019], FastMICE [Huang *et al.*, 2023], MvDGNMF [Li *et al.*, 2020a], MVC-DMF-PA [Zhang *et al.*, 2021], AGFS-OMVC [Zhao *et al.*, 2024], Orth-NTF [Li *et al.*, 2024b], MVC-DNTF [Feng *et al.*, 2024].

4.3 Experimental Setup

All experiments were executed on a desktop with an Intel(R) Core(TM) i5-13400 CPU and 32 GB of RAM, using MATLAB 2023a. Data normalization was performed as a preprocessing step for all datasets to ensure consistent input quality. We assessed the clustering quality using Accuracy (ACC), Normalized Mutual Information (NMI), and Purity (PUR).

For the deep anchor graph projection approach, the layer size configurations were determined based on the complexity of the dataset. Specifically, for a two-layer projection, sizes $[d_1, k]$ were used, where d_1 varied among $[4k, 5k, 6k]$. For a three-layer configuration, sizes $[d_1, d_2, k]$ were used, with d_1 in the range of $[8k, 10k, 12k]$ and d_2 between $[4k, 5k, 6k]$. These configurations allowed us to explore the impact of depth in the deep projection on the clustering results.

4.4 Clustering Performance

The experimental results presented in Tables 2 and 3 demonstrate the clustering performance of our proposed method cross multiple datasets in comparison with several benchmark algorithms. The best results are highlighted in bold, while the second-best results are underlined.

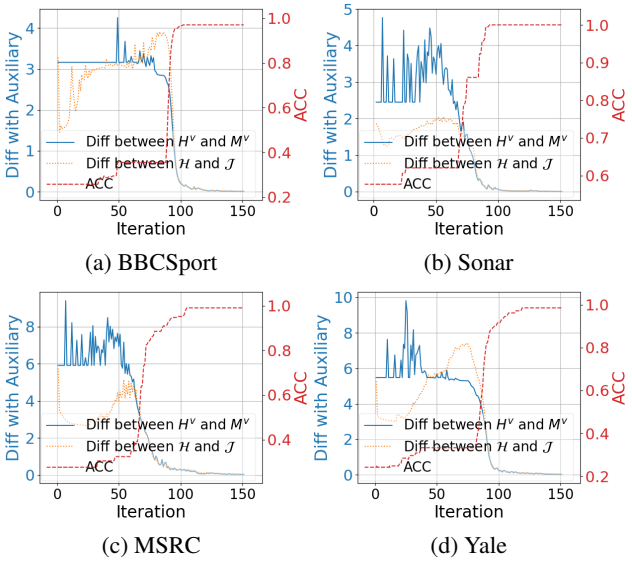
The results indicate that the deep projection structure in TMVC-DAGP significantly enhances clustering accuracy by capturing complex data relationships more effectively. Com-

pared with AGFS-OMVC, which also employs a projection-based approach, and Orth-NTF, which leverages tensor decomposition, TMVC-DAGP consistently outperforms them. For instance, on the RGB-D dataset, TMVC-DAGP achieved an ACC of 72.25%, surpassing the best-performing baselines. This highlights the advantage of our deep structure in extracting meaningful information from multi-view data.

Furthermore, the direct projection of anchor graphs into the label space contributes to the robustness and precision of TMVC-DAGP. In contrast to MVC-DMF-PA and MVC-DNTF, which utilize deep non-negative matrix factorization, our approach demonstrates superior performance. Notably, TMVC-DAGP achieved an ACC of 98.78% on the Yale dataset and superior clustering results to other methods aon the Sonar and Vehicle Sensor datasets. The integration of tensor-based view fusion further strengthens clustering performance by effectively leveraging complementary information across different views.

4.5 Convergence

To validate the convergence of TMVC-DAGP, we conducted comprehensive experiments on multiple datasets by tracking the changes in objective function values, specifically focusing on the differences between $\mathbf{H}^v$ and $\mathbf{M}^v$, as well as $\mathcal{H}$ and $\mathcal{J}$. The convergence curves for the BBCSport, Sonar, MSRC, and Yale datasets are presented in Figure 1. The results demonstrate a consistent decrease in objective function values as iterations increase, reflecting the stability and robustness of the method. For instance, the BBCSport dataset achieves convergence after approximately 100 iterations, while Sonar, MSRC, and Yale datasets converge within 80 iterations. This rapid and consistent convergence across diverse datasets underscores the effectiveness of TMVC-DAGP in practice. This is a sentence in 7pt font size.

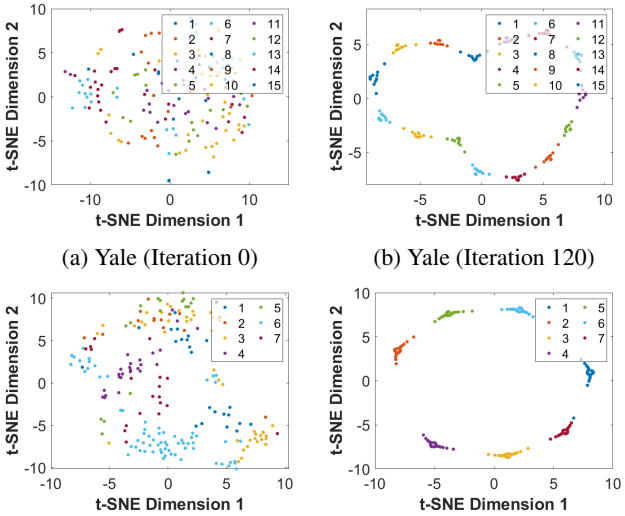


(a) BBCSport (b) Sonar

(c) MSRC (d) Yale

Figure 1: Convergence curve on BBCSport, Sonar, MSRC, and Yale

Additionally, to validate the clustering process and evaluate how well the data structures align during iterations, we employed t-SNE visualizations. These visualizations demonstrate the evolution of cluster separability at different stages of the iterative process, as shown in Figure 2.



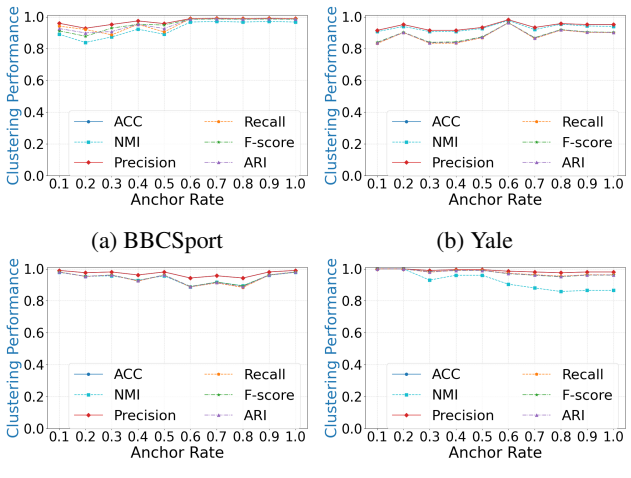
(a) Yale (Iteration 0) (b) Yale (Iteration 120)

(c) MSRC (Iteration 0) (d) MSRC (Iteration 80)

Figure 2: t-SNE visualizations of the BBCSport, NGS, RGB-D, and MSRC datasets at different iterations.

Overall, the convergence analysis and visualizations confirm that TMVC-DAGP achieves both numerical stability and enhanced clustering performance in multi-view scenarios.

4.6 Parameter Analysis

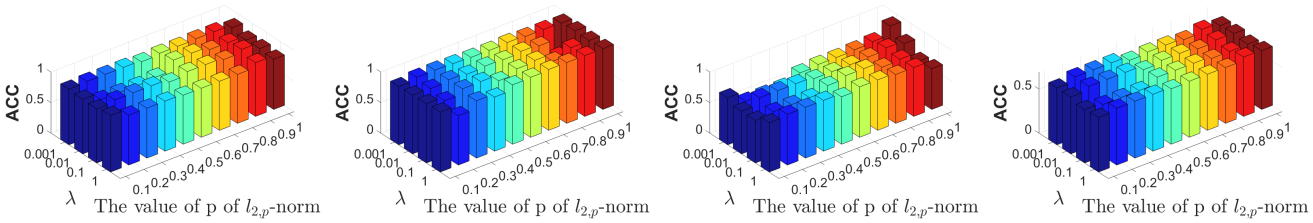


(a) BBCSport (b) Yale

(c) MSRC (d) sonar

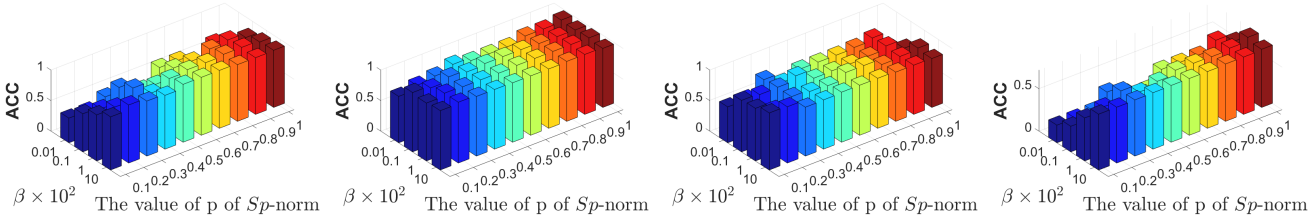
Figure 3: Clustering results with different anchor rates on 3-sources, BBCSport, Yale, RGB-D

We study how different parameters affect the clustering results. First, we analyze the impact of the anchor rate on clustering performance. Figure 3 illustrates ACC under different anchor rates on BBCSport, Yale, MSRC, and Sonar datasets. It can be observed that choosing an appropriate anchor rate is



(a) BBCSport (b) Sonar (c) Yale (d) RGB-D

Figure 4: The influence of λ and p (belonging to the $\ell_{2,p}$-norm) on clustering results for the BBCSport, Sonar, Yale, and RGB-D.



(a) BBCSport (b) Sonar (c) Yale (d) RGB-D

Figure 5: The influence of β and p (belonging to the Schatten p-norm) on clustering results for the BBCSport, Sonar, Yale, and RGB-D.

important for improving clustering performance, as too high or too low anchor rates may lead to decreased performance.

Next, we explore the effects of the regularization parameters λ, β, p of $\ell_{2,p}$ norm and p of Schatten p norm in conjunction with clustering performance. Figures 4 and 5 illustrate the impact of different combinations of these parameters across several datasets. This demonstrates the robustness of TMVC-DAGP, as it maintains effective clustering performance even when parameters are varied.

4.7 Ablation Study

We conducted ablation experiments to evaluate the effectiveness of each component. First, we compared the performance of the algorithm when the Schatten p-norm (case1) and $\ell_{2,p}$-norm (case2) constraints were retained individually. Table 4 shows the clustering accuracy under different settings across four benchmark datasets. The results indicate that removing any of these constraints leads to a drop in performance, especially on complex datasets, showing that these constraints are crucial for maintaining stability and improving accuracy.

case1	case2	BBCSport	Yale	RGB-D	Sonar
×	×	67.83	50.31	55.76	60.09
×	✓	88.78	85.45	62.94	93.26
✓	×	92.46	91.51	70.87	97.11
✓	✓	99.08	98.78	72.25	100.00

Table 4: ACC(%) of ablation experiments

We also studied the effect of the layer number in the deep projection structure, as shown in Table 5 and Table 6. It demonstrates that increasing the layer number can effectively improve clustering performance but will not introduce much computational burden, as deeper structures better capture the intrinsic structure of the data. Considering the improvement of performance, the extra computational cost is acceptable.

p	BBCSport	Yale	RGB-D	Sonar
$[k]$	97.06	97.40	68.32	99.03
$[d_1, k]$	98.05	98.18	71.01	99.03
$[d_1, d_2, k]$	99.08	98.18	72.25	100.00

Table 5: ACC(%) of different layers on four datasets.

p	BBCSport	Yale	RGB-D	Sonar
$[k]$	3.10	1.26	16.00	1.24
$[d_1, k]$	4.58	1.54	17.16	1.23
$[d_1, d_2, k]$	4.45	1.78	18.85	1.32

Table 6: Running time (s) of different layers on four datasets.

5 Conclusion

In this paper, we introduced a novel Tensorial Multi-view Clustering with Deep Anchor Graph Projection (TMVC-DAGP), which effectively captures the complex structures via projecting the anchor graph into label space. Together with tensorized schatten p-norm, TMVC-DAGP exploits complementary information across various views. Besides, a sparsity constraint using the $\ell_{2,p}$-norm enhances robustness and achieves one-step clustering, particularly in high-dimensional and noisy environments. Extensive experiments and comparisons on several multi-view datasets illustrate the effectiveness and superiority of our method.

Acknowledgments

This work is supported by This work is supported by the National Science Foundation of China (No. 62037001), the Key Research and Development Project in Shaanxi Province (2024PT-ZCK-89), the Fundamental Research Funds for the Central Universities (ZYTS25267, QTZX25004), and the Science and Technology Project of Xi'an (Grant 2022JH-JSYF-0009), Open Project of Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, Anhui University (No. MMC202416), Selected Support Project for Scientific and Technological Activities of Returned Overseas Chinese Scholars in Shaanxi Province 2023-02, and the Xidian Innovation Fund (Project NoYJSJ25007).

References

[Blum and Mitchell, 1998] Avrim Blum and Tom Mitchell. Combining labeled and unlabeled data with co-training. In *Proceedings of the eleventh annual conference on Computational learning theory*, pages 92–100, 1998.

[Brbić and Kopriva, 2018] Maria Brbić and Ivica Kopriva. Multi-view low-rank sparse subspace clustering. *Pattern Recognition*, 73:247–258, 2018.

[Cao *et al.*, 2015] Xiaochun Cao, Changqing Zhang, Huazhu Fu, Si Liu, and Hua Zhang. Diversity-induced multi-view subspace clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–594, 2015.

[Chen *et al.*, 2022] Man-Sheng Chen, Chang-Dong Wang, and Jian-Huang Lai. Low-rank tensor based proximity learning for multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 35(5):5076–5090, 2022.

[Duarte and Hu, 2004] Marco F Duarte and Yu Hen Hu. Vehicle classification in distributed sensor networks. *Journal of Parallel and Distributed Computing*, 64(7):826–838, 2004.

[Fang *et al.*, 2023] Si-Guo Fang, Dong Huang, Xiao-Sha Cai, Chang-Dong Wang, Chaobo He, and Yong Tang. Efficient multi-view clustering via unified and discrete bipartite graph learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[Feng *et al.*, 2024] Wei Feng, Dongyuan Wei, Qianqian Wang, Bo Dong, and Quanxue Gao. Multi-view clustering based on deep non-negative tensor factorization. In *ACM Multimedia 2024*, 2024.

[Fu *et al.*, 2020] Lele Fu, Pengfei Lin, Athanasios V Vasilakos, and Shiping Wang. An overview of recent multi-view clustering. *Neurocomputing*, 402:148–161, 2020.

[Gao *et al.*, 2020a] Quanxue Gao, Zhizhen Wan, Ying Liang, Qianqian Wang, Yang Liu, and Ling Shao. Multi-view projected clustering with graph learning. *Neural Networks*, 126:335–346, 2020.

[Gao *et al.*, 2020b] Quanxue Gao, Wei Xia, Zhizhen Wan, Deyan Xie, and Pu Zhang. Tensor-svd based graph learning for multi-view subspace clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 3930–3937, 2020.

[Greene and Cunningham, 2006] Derek Greene and Pádraig Cunningham. Practical solutions to the problem of diagonal dominance in kernel document clustering. In *Proceedings of the 23rd international conference on Machine learning*, pages 377–384, 2006.

[Guo *et al.*, 2022] Jipeng Guo, Yanfeng Sun, Junbin Gao, Yongli Hu, and Baocai Yin. Logarithmic schatten-p p norm minimization for tensorial multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(3):3396–3410, 2022.

[Huang *et al.*, 2023] Dong Huang, Chang-Dong Wang, and Jian-Huang Lai. Fast multi-view clustering via ensembles: Towards scalability, superiority, and simplicity. *IEEE Transactions on Knowledge and Data Engineering*, 2023.

[Hussain *et al.*, 2010] Syed Fawad Hussain, Gilles Bisson, and Clément Grimal. An improved co-similarity measure for document clustering. In *2010 ninth international conference on machine learning and applications*, pages 190–197. IEEE, 2010.

[Jiang *et al.*, 2013] Yu Jiang, Jing Liu, Zechao Li, Peng Li, and Hanqing Lu. Co-regularized plsa for multi-view clustering. In *Computer Vision–ACCV 2012: 11th Asian Conference on Computer Vision, Daejeon, Korea, November 5-9, 2012, Revised Selected Papers, Part II 11*, pages 202–213. Springer, 2013.

[Lei *et al.*, 2024] Yu Lei, Zuoyuan Niu, Qianqian Wang, Quanxue Gao, and Ming Yang. Anchor graph-based multiview spectral clustering. *Neurocomputing*, 583:127579, 2024.

[Li *et al.*, 2015] Yeqing Li, Feiping Nie, Heng Huang, and Junzhou Huang. Large-scale multi-view spectral clustering via bipartite graph. In *Proceedings of the AAAI conference on artificial intelligence*, volume 29, 2015.

[Li *et al.*, 2019] Ruihuang Li, Changqing Zhang, Huazhu Fu, Xi Peng, Tianyi Zhou, and Qinghua Hu. Reciprocal multi-layer subspace learning for multi-view clustering. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 8172–8180, 2019.

[Li *et al.*, 2020a] Jianqiang Li, Guoxu Zhou, Yuning Qiu, Yanjiao Wang, Yu Zhang, and Shengli Xie. Deep graph regularized non-negative matrix factorization for multi-view clustering. *Neurocomputing*, 390:108–116, 2020.

[Li *et al.*, 2020b] Xuelong Li, Han Zhang, Rong Wang, and Feiping Nie. Multiview clustering: A scalable and parameter-free bipartite graph fusion method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1):330–344, 2020.

[Li *et al.*, 2023] Jing Li, Qianqian Wang, Ming Yang, Quanxue Gao, and Xinbo Gao. Efficient anchor graph factorization for multi-view clustering. *IEEE Transactions on Multimedia*, 2023.

[Li *et al.*, 2024a] Jing Li, Quanxue Gao, Qianqian Wang, Cheng Deng, and Deyan Xie. Label learning method based

on tensor projection. *arXiv preprint arXiv:2402.16544*, 2024.

[Li *et al.*, 2024b] Jing Li, Quanxue Gao, Qianqian Wang, Ming Yang, and Wei Xia. Orthogonal non-negative tensor factorization based multi-view clustering. *Advances in Neural Information Processing Systems*, 36, 2024.

[Li *et al.*, 2024c] Xingfeng Li, Yuangang Pan, Yuan Sun, Quansen Sun, Yinghui Sun, Ivor W Tsang, and Zhenwen Ren. Incomplete multi-view clustering with paired and balanced dynamic anchor learning. *IEEE Transactions on Multimedia*, 2024.

[Liu *et al.*, 2021] Jiyuan Liu, Xinwang Liu, Yuexiang Yang, Xifeng Guo, Marius Kloft, and Liangzhong He. Multiview subspace clustering via co-training robust data representation. *IEEE Transactions on Neural Networks and Learning Systems*, 33(10):5177–5189, 2021.

[Qin *et al.*, 2024] Yalan Qin, Chuan Qin, Xinpeng Zhang, and Guorui Feng. Dual consensus anchor learning for fast multi-view clustering. *IEEE Transactions on Image Processing*, 2024.

[Sejnowski and Gorman,] Terry Sejnowski and R Gorman. Connectionist Bench (Sonar, Mines vs. Rocks). UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C5T01Q.

[Silberman *et al.*, 2012] Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. Indoor segmentation and support inference from rgbd images. In *Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part V 12*, pages 746–760. Springer, 2012.

[Tzortzis and Likas, 2012] Grigorios Tzortzis and Aristidis Likas. Kernel-based weighted multi-view clustering. In *2012 IEEE 12th international conference on data mining*, pages 675–684. IEEE, 2012.

[Wang *et al.*, 2019] Hao Wang, Yan Yang, and Bing Liu. Gmc: Graph-based multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 32(6):1116–1129, 2019.

[Wang *et al.*, 2021a] Qi Wang, Ran Liu, Mulin Chen, and Xuelong Li. Robust rank-constrained sparse learning: A graph-based framework for single view and multi-view clustering. *IEEE Transactions on Cybernetics*, 52(10):10228–10239, 2021.

[Wang *et al.*, 2021b] Siwei Wang, Xinwang Liu, Xinzhong Zhu, Pei Zhang, Yi Zhang, Feng Gao, and En Zhu. Fast parameter-free multi-view subspace clustering with consensus anchor guidance. *IEEE Transactions on Image Processing*, 31:556–568, 2021.

[Wei *et al.*, 2017] Xiaokai Wei, Bokai Cao, and S Yu Philip. Multi-view unsupervised feature selection by cross-diffused matrix alignment. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 494–501. IEEE, 2017.

[Winn and Jojic, 2005] John Winn and Nebojsa Jojic. Locus: Learning object classes with unsupervised segmentation. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 1, pages 756–763. IEEE, 2005.

[Xia *et al.*, 2021] Wei Xia, Xiangdong Zhang, Quanxue Gao, Xiaochuang Shu, Jungong Han, and Xinbo Gao. Multiview subspace clustering by an enhanced tensor nuclear norm. *IEEE Transactions on cybernetics*, 52(9):8962–8975, 2021.

[Yale University, 2001] Yale University. yale. http://cvc.cs.yale.edu/cvc/projects/yalefaces/yalefaces.html, 2001.

[Yang *et al.*, 2022] Ben Yang, Xuetao Zhang, Zhongheng Li, Feiping Nie, and Fei Wang. Efficient multi-view k-means clustering with multiple anchor graphs. *IEEE Transactions on Knowledge and Data Engineering*, 35(7):6887–6900, 2022.

[Zhang *et al.*, 2021] Chen Zhang, Siwei Wang, Jiyuan Liu, Sihang Zhou, Pei Zhang, Xinwang Liu, En Zhu, and Changwang Zhang. Multi-view clustering via deep matrix factorization and partition alignment. In *Proceedings of the 29th ACM international conference on multimedia*, pages 4156–4164, 2021.

[Zhao *et al.*, 2017] Handong Zhao, Zhengming Ding, and Yun Fu. Multi-view clustering via deep matrix factorization. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31, 2017.

[Zhao *et al.*, 2024] Wenhui Zhao, Qin Li, Huafu Xu, Quanxue Gao, Qianqian Wang, and Xinbo Gao. Anchor graph-based feature selection for one-step multi-view clustering. *IEEE Transactions on Multimedia*, 2024.

[Zheng *et al.*, 2023] Xiao Zheng, Chang Tang, Xinwang Liu, and En Zhu. Multi-view clustering via matrix factorization assisted k-means. *Neurocomputing*, 534:45–54, 2023.