PRELIMINARY PREPRINT VERSION – DO NOT CITE
This is not the version of record. The final version will appear in the Proceedings of IJCAI-ECAI 2026.

Cooperative Multi-View Graph Learning via High-Rank Tensor Specificity

Ling Ma [1] , **Qiyu Zhong**[1] , **Songxuan Shi**[1] , **Hao Wei**[1] , **Shunjie Yang**[1] , **Junbo Lian**[2] , **Qiuru Hai**[1] , **Lianjin Yu**[1] , **Xiangning Zeng**[1] , **Yi Shan**[1] , **Zhen Yang**[1] , **Gengyu Lyu**[1,*]
[1]Beijing University of Technology
[2]Northwestern University
maling@emails.bjut.edu.cn, zhongqy_ovo@163.com, shisongxuan@emails.bjut.edu.cn,
haowei@emails.bjut.edu.cn, shunjieyang@163.com, jacoblian@u.northwestern.edu,
haiqiuru@emails.bjut.edu.cn, yulianjin20041212@gmail.com, zengxiangning2025@163.com,
yi.shan.bjut@gmail.com, yangzhen@bjut.edu.cn, lyugengyu@gmail.com

Abstract

Graph-based multi-view clustering, with its ability to mine potential associations between samples, has attracted extensive attention. To capture high-order correlations, tensor-based frameworks have been introduced to model multiple graphs jointly. Although these methods have achieved promising performance, existing methods mainly focus on stacking consistency graphs with low-rank constraints, while overlooking high-order specificity within diversity graphs, and the exploration of fine-grained diversity information among different samples across views remains insufficient. To address these issues, we propose **H**igh-Rank **T**ensor **S**pecificity Induced **C**ooperative **M**ulti-View **G**raph **L**earning (HTS-CMGL). Specifically, we first obtain the consistency graphs and diversity graphs from multi-view data, which are further reconstructed into the consistency tensor and the diversity tensor, respectively. Subsequently, a novel Enhanced Tensor Rank (ETR) is imposed on the consistency tensor, which is a tighter approximation of the tensor rank and is more noisy-robust to explore the high-order consistency. Meanwhile, we design a new Tensor High-Rank Logarithmic Norm (THLN) on the diversity tensor. By leveraging a unique high-rank constraint mechanism, THLN not only actively preserves high-rank and informative features, but also simultaneously captures view-level and sample-level diversity information. Extensive experiments on multiple datasets demonstrate the effectiveness of our proposed method on clustering multi-view data.

1 Introduction

Multi-View Clustering (MVC) has emerged as a dominant paradigm for integrating heterogeneous information from diverse sources to uncover data partitions [Yang *et al.*, 2025; Yang *et al.*, 2026]. Graph learning [Li *et al.*, 2026; Lin *et al.*, 2025; Lin *et al.*, 2026] has become a promising technique for multi-view clustering due to its efficiency in learning a unified graph from multiple views. Early approaches primarily focused on exploiting multi-view consistency [Zhan *et al.*, 2019]. Subsequently, representative studies such as [Liang *et al.*, 2019] and [Huang *et al.*, 2022] have explicitly modeled the non-consensus components (termed diversity) within multi-view data. These studies emphasize the need to explicitly model diversity from noise, corruptions, or view-specific attributes. Such separation effectively removes noise interference, yielding a more robust consensus graph. This philosophy also proves effective in federated learning [Sun *et al.*, 2025a; Sun *et al.*, 2025b].

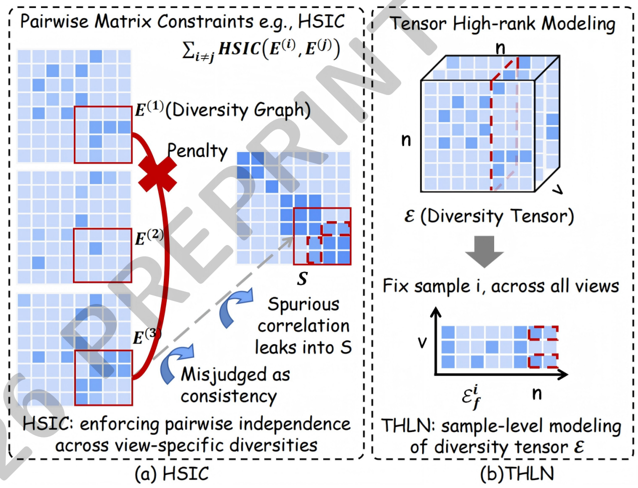


Figure 1: Motivation illustration. (a) HSIC imposes global matrix-level pairwise constraints, considering only overall inter-view diversity and potentially misjudging local co-occurrences. (b) THLN tensorizes diversity and models it sample-wise, capturing high-order specificity and its heterogeneous distribution across samples.

However, such methods are limited to pairwise interactions and fail to capture high-order structures. Consequently, tensor-based algorithms [Xia *et al.*, 2023; Tang *et al.*, 2025; Feng *et al.*, 2025], stack consistency matrices into a representation tensor and impose low-rank constraints to capture high-order correlations. For instance, RTGD-MVC [Zou *et al.*, 2025] utilizes Tensor Exponential Norm (TEN) minimization to capture high-order correlations, while suppressing view-specific inconsistencies via $\ell_{2,1}$-norm constraints. DETDSI [Sun *et al.*, 2025c] introduces Enhanced Tensor Rank (ETR) minimization to capture high-order correlations, and employs the Hilbert-Schmidt Independence Criterion (HSIC [Lyu *et al.*, 2024]) to enforce statistical independence on the extracted inconsistencies. Despite achieving competitive performance, these methods face two challenges: (1) On one hand, existing methods rely on matrix-level regularization to handle view-specific diversity (e.g., HSIC-based dependence penalties or $\ell_{2,1}$-norm robustness terms). (2) On the other hand, existing methods treat inter-view diversity as a global quantity shared by all samples, whereas in practice the degree of inter-view diversity varies across samples and follows a heterogeneous distribution. As illustrated in Figure 1(a), HSIC treats inter-view diversity as a global property, which may suppress locally co-occurring, sample-specific diversity and blur the boundary between diversity and consistency.

To mitigate these problems, we propose a novel tensorized MVC method termed High-Rank Tensor Specificity Induced Cooperative Multi-View Graph Learning (HTS-CMGL). As shown in Figure 1(b), HTS-CMGL stacks multi-view diversity graphs into a tensor and performs sample-wise slice modeling to characterize heterogeneous diversity across samples and capture high-order specificity. To our knowledge, this is the first attempt to explicitly model diversity as a high-rank tensor, thereby capturing fine-grained structural details that are suppressed by traditional pairwise matrix constraints. Specifically, we first decompose multi-view similarity graphs into consistency graphs and diversity graphs, reconstructing them into tensors to exploit high-order consistency and high-order diversity. Subsequently, an Enhanced Tensor Rank (ETR) constraint is imposed on the consistency tensor, which preserves larger singular values and discards smaller ones to obtain an accurate low-rank tensor representation. Meanwhile, we propose the Tensor High-Rank Logarithmic Norm (THLN), which enables the exploration of high-order cross-view specificity while modeling diversity across samples and views. Moreover, owing to its intrinsic high-rank property, THLN allows the model to actively learn rich and structured view-specific diversity. This adversarial separation purifies both the consistency tensor and the final consensus graph. Extensive experimental results on multiple benchmark datasets demonstrate that our HTS-CMGL exhibits superior performance compared to other state-of-the-art methods. The main contributions of our paper are summarized as:

- We propose HTS-CMGL, a tensorized graph learning method that establishes an adversarial separation mechanism between consistency and diversity. By simultaneously imposing the ETR for high-order consistency and the THLN as a high-rank constraint on diversity, our model achieves a more thorough separation between consistency and diversity.
- To better capture high-order specificity, we propose a novel Tensor High-Rank Logarithmic Norm (THLN), which actively preserves the intrinsic high-rank structure of diversity and characterizes the heterogeneous distribution of diversity across samples.
- Extensive comparative experiments and detailed analysis have shown that our proposed HTS-CMGL method can achieve significantly superior performance against other state-of-the-art methods.

2 Related Work

2.1 Tensor-based Multi-View Clustering

Tensor-based methods are designed to exploit the high-order correlations among views [Bao *et al.*, 2025; Zhen *et al.*, 2025; Guo *et al.*, 2024; Wang *et al.*, 2025]. Tensor nuclear norm minimization (TNN), particularly when combined with tensor singular value decomposition (t-SVD), has emerged as a powerful tool to model high-order correlations and preserve the spatial structure of multi-view data [Xie *et al.*, 2018]. For instance, [Tang *et al.*, 2025] proposed high-order bipartite graph learning and tensor low-rank representation (HBGTLRR), which concatenates high-order bipartite graphs into a tensor and imposes the TNN constraint, effectively reducing noise while preserving high-order consistency. However, TNN is a loose and biased approximation of the true tensor rank, as it treats different singular values equally, leading to residual noise and loss of important information. To address this problem, several non-convex approximations have been proposed, such as the Tensor Logarithmic Schatten-p Norm (TLS_pN) [Guo *et al.*, 2023], the Enhanced Tensor Rank (ETR) [Ji and Feng, 2025] and the Laplace Tensor Rank (LTR) [Zhong *et al.*, 2025]. Despite these advancements, methods predominantly focus on imposing low-rank constraints on consistent tensors, overlooking the potential of tensors to explicitly capture complex high-order specificity.

2.2 Graph-based Multi-view Clustering

Graph-based multi-view clustering methods have demonstrated satisfying performance by effectively capturing relationships among samples. These methods primarily focus on enhancing the quality of the consensus graph and improving the purity of the view-specific similarity graphs to optimize clustering performance [Peng *et al.*, 2025; Sun *et al.*, 2025b; **?**]. [Zhao *et al.*, 2025] introduced adaptive weighting mechanisms to assign optimal contributions, ensuring that high-quality views dominate the fusion process. Similarly, [Huang *et al.*, 2024] separately explores the consistent information and complementary information by dividing the affinity matrix in each view into two parts. To explicitly capture high-order correlations across multi views, tensor-based frameworks have been introduced. For example, [Xia *et al.*, 2023] proposed TBGL, which utilizes Tensor Schatten p-norm minimization to capture high-order consistency, while suppressing view-specific inconsistencies via ℓ_1-norm constraints. However, the aforementioned methods typically rely on HSIC or simple norm constraints to decouple diversity. These approaches are limited to modeling pairwise view relationships and implicitly assume that view discrepancies are stationary across all samples. Consequently, they lack a mechanism to explicitly capture cross-view, sample-level specificity.

3 The Proposed Method

3.1 Motivations and Problem Statement

Graph-based multi-view clustering has attracted significant attention due to its capability to uncover latent correlations among data samples. However, current methods generally confront the following challenges: (1) **oversimplification**

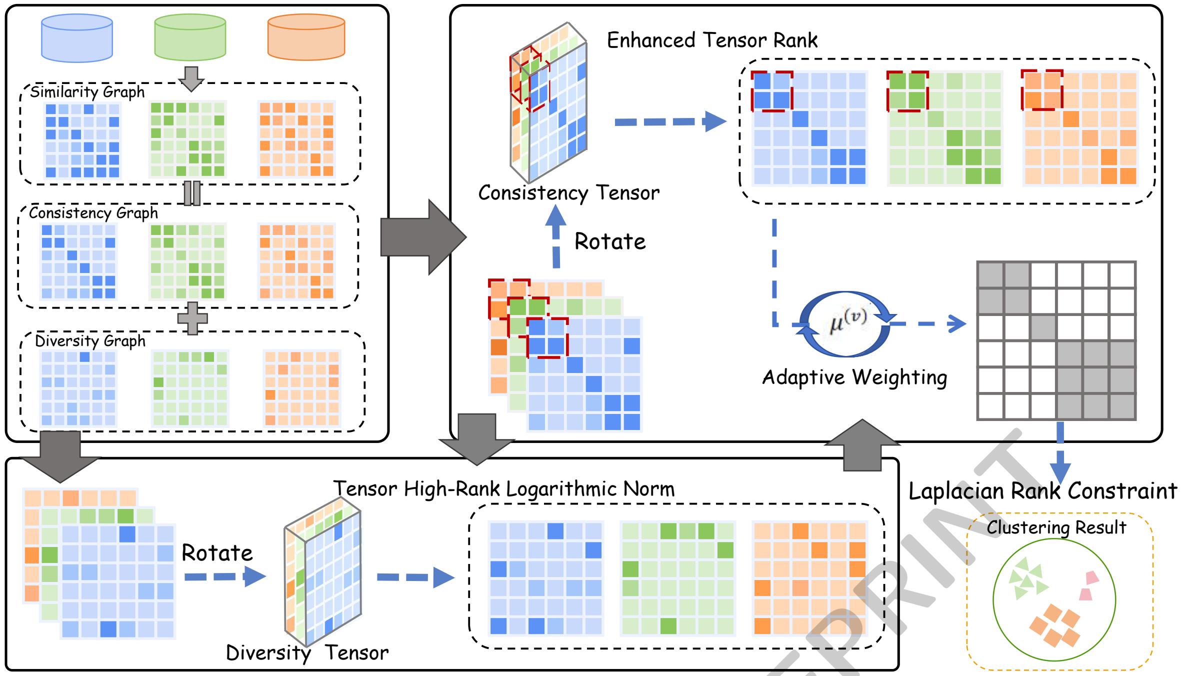


Figure 2: The framework of our HTS-CMGL. We first construct the similarity graphs for each view. Then, we introduce the ETR and THLN to capture the high-order correlations and high-order specificity, respectively. Finally, a weighted fusion strategy is employed to obtain the consensus graph for spectral clustering.

of diversity distribution, where diversity is assumed to be stationary across samples despite being heterogeneous and sample-specific; (2) **matrix-level modeling of view-specific diversity**, which neglects the exploration of high-order specificity, and where matrix-level constraints may suppress locally co-occurring diversity; and (3) **biased approximation of low-rank structure**, TNN treats all singular values equally, which often leads to insufficient suppression of noise-related components.

To effectively address these limitations, we propose a novel framework termed HTS-CMGL. Specifically, we introduce the ETR to rectify the biased approximation of the low-rank structure, and design a novel THLN to actively explore fine-grained diversity information among different samples across views in a unified and cooperative manner.

Problem Formulation

In this paper, scalars, vectors, matrices, and tensors are denoted by lowercase letters x, bold lowercase letters $\mathbf{x}$, bold uppercase letters $\mathbf{X}$, and calligraphic letters $\mathcal{X}$, respectively. For a 3-way tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, $\mathcal{X}^{(k)}$ denotes the k-th frontal slice, and $\mathcal{X}_f$ represents its Fast Fourier Transform (FFT) along the third dimension.

We consider a multi-view dataset $\{\mathbf{X}^{(v)}\}_{v=1}^{V}$, where $\mathbf{X}^{(v)} \in \mathbb{R}^{n \times d_v}$ contains n samples with d_v features in the v-th view. For each view, we construct an affinity matrix $\mathbf{A}^{(v)}$ to model pairwise similarities among samples. Since raw graphs often mix shared structures with view-specific variations and noise, each $\mathbf{A}^{(v)}$ is explicitly decomposed into a consistent component $\mathbf{C}^{(v)}$ and a diversity component $\mathbf{E}^{(v)}$, i.e., $\mathbf{A}^{(v)} = \mathbf{C}^{(v)} + \mathbf{E}^{(v)}$.

Adaptive Consensus Graph Learning

To learn a unified graph $\mathbf{S}$ with a clear clustering structure, we minimize its discrepancy from the view-specific consistent representations $\{\mathbf{C}^{(v)}\}_{v=1}^{V}$ and assign weights based on the reliability and information volume of different views. Since uniform weight allocation treats all views equally and leaves the consensus graph susceptible to noise, we adopt an adaptive weight scheme to balance the contributions of different views. Furthermore, we introduce a Laplacian rank constraint to explicitly impose a c-cluster structure on $\mathbf{S}$. The resulting optimization problem is formulated as:

$$\begin{aligned} \min_{\mathbf{S},\mathbf{F},\boldsymbol{\mu}} \quad & \sum_{v=1}^{V} (\mu^{(v)})^2 \left\| \mathbf{S} - \mathbf{C}^{(v)} \right\|_{\mathrm{F}}^2 + 2\lambda \mathrm{Tr}(\mathbf{F}^\top \mathbf{L}_{\mathbf{S}} \mathbf{F}) \\ \text{s.t.} \quad & \mathbf{s}_i^\top \mathbf{1} = 1,\ \mathbf{S} \geq 0,\ \mathbf{F}^\top \mathbf{F} = \mathbf{I}_c,\ \boldsymbol{\mu}^\top \mathbf{1} = 1, \end{aligned} \tag{1}$$

where $\mu^{(v)}$ denotes the adaptive weight of the v-th view; unlike a linear weighting scheme, the squared weighting yields a non-degenerate closed-form update under the simplex constraint and thus avoids the trivial *winner-take-all* solution.

Here, $\mathbf{L}_{\mathbf{S}} = \mathbf{D}_{\mathbf{S}} - \frac{\mathbf{S}+\mathbf{S}^\top}{2}$ denotes the graph Laplacian. By Ky Fan's theorem [Fan, 1949], minimizing $\mathrm{Tr}(\mathbf{F}^\top \mathbf{L}_{\mathbf{S}} \mathbf{F})$ leads to $\mathrm{rank}(\mathbf{L}_{\mathbf{S}}) = N - c$, which implies that the similarity graph $\mathbf{S}$ consists of c connected components.

The optimal weights admit a closed-form solution:

$$\mu^{(v)} = \frac{1/\left\| S - C^{(v)} \right\|_F^2}{\sum_v 1/\left\| S - C^{(v)} \right\|_F^2} \tag{2}$$

Tensorized Consistency and Specificity Learning

While graph-level fusion handles pairwise relations, it fails to capture the high-order correlations and high-order specificity embedded across multiple views. To bridge this gap, we propose a unified Dual-Rank Constraint Strategy. We tensorize the consistency graphs $\{\mathbf{C}^{(v)}\}$ and diversity graphs $\{\mathbf{E}^{(v)}\}$ into $\mathcal{C}$ and $\mathcal{E}$ via a rotation operator $\Phi(\cdot)$. This structure aligns sample-wise similarities across views into tensor tubes, thereby enabling the simultaneous exploration of high-order consensus and high-order diversity via tensor decomposition.

1) Consistency via Low-Rank:

Definition 1. Given a tensor $\mathcal{C} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the Enhanced Tensor Rank (ETR)[Ji and Feng, 2025] is defined as:

$$\begin{aligned} \|\mathcal{C}\|_{\mathrm{ETR}} &= \frac{1}{n_3} \sum_{k=1}^{n_3} \|\mathcal{C}_f^k\|_{\mathrm{ETR}} \\ &= \frac{1}{n_3} \sum_{k=1}^{n_3} \sum_{i=1}^{h} \left(\frac{e^{\delta^2} \mathcal{S}_f^k(i,i)}{\delta + \mathcal{S}_f^k(i,i)} \right), \end{aligned} \tag{3}$$

where $0 < \delta \leq 1$, $h = \min(n_1, n_2)$, and $\mathcal{S}_f$ is obtained by t-SVD of $\mathcal{C}_f = \mathcal{U}_f \mathcal{S}_f \mathcal{V}_f^\top$ in the Fourier domain.

To explore the high-order consistency shared across multiple views, we impose a low-rank constraint on the consistency tensor $\mathcal{C}$. Instead of adopting the conventional TNN, which is known to be a loose and biased approximation of tensor rank, we employ a more noise-robust ETR to more accurately characterize the intrinsic low-rank structure of $\mathcal{C}$.

2) Specificity via High-Rank Preservation:

Definition 2. Given a tensor $\mathcal{E} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the Tensor High-Rank Logarithmic Norm (THLN) is defined as:

$$\begin{aligned} \|\mathcal{E}\|_{\mathrm{THLN}} &= \frac{1}{n_3} \sum_{k=1}^{n_3} \|\mathcal{E}_f^k\|_{\mathrm{THLN}} \\ &= \frac{1}{n_3} \sum_{k=1}^{n_3} \sum_{i=1}^{h} \left(-\mathcal{S}_f^k(i,i) \log(1 + \mathcal{S}_f^k(i,i)) \right), \end{aligned} \tag{4}$$

where $h = \min(n_1, n_2)$ and $\mathcal{S}_f$ is obtained by t-SVD of $\mathcal{E}_f = \mathcal{U}_f \mathcal{S}_f \mathcal{V}_f^\top$ in the Fourier domain.

Remark [The Superiority of THLN]: From a statistical perspective, pairwise matrix-level dependence modeling is insufficient to characterize high-order co-occurrence patterns among multiple variables [Pfister *et al.*, 2018]. Different from matrix-level approaches that only explore pairwise complementarity, our THLN is a tensor-level high-rank constraint that captures high-order specificity by modeling all views holistically. Furthermore, while existing methods implicitly assume that the diversity between views is stationary (uniformly distributed) across all samples, we recognize that diversity is inherently **heterogeneous and sample-specific**. THLN actively encourages complex and incompressible structures, effectively maintaining informative diversity patterns that capture these fine-grained variations.

By integrating adaptive consensus graph learning with tensorized consistency and specificity modeling, we formulate the final optimization problem as

$$\begin{aligned} &\min_{\mathrm{S,\,C}} \sum_{v=1}^{V} (\mu^{(v)})^2 \left\| \mathbf{S} - \mathbf{C}^{(v)} \right\|_{\mathrm{F}}^2 + \beta \|\mathcal{C}\|_{\mathrm{ETR}} + \gamma \|\mathcal{E}\|_{\mathrm{THLN}} \\ &\quad + 2\lambda \mathrm{Tr}(F^T L_S F) \\ &\mathcal{C} = \Phi\left(\mathbf{C}^{(1)}, \cdots, \mathbf{C}^{(v)}\right), \mathcal{E} = \Phi\left(\mathbf{E}^{(1)}, \cdots, \mathbf{E}^{(v)}\right), \\ &s.t. \forall v, s^t 1 = 1, \mathbf{S} \geq 0, \mathcal{A} = \mathcal{C} + \mathcal{E},\ \mathbf{A}^{(v)} \geq \mathbf{C}^{(v)} \geq 0, \end{aligned} \tag{5}$$

where β and γ are the trade-off parameters.

3.2 Optimization

Inspired by the alternating direction method of multipliers (ADMM) [Lin *et al.*, 2011], we introduce the auxiliary tensor variables $\mathcal{J}, \mathcal{G}$, so the model Eq.(5) can be rewritten as the following unconstrained problem,

$$\begin{aligned} &\mathcal{L}(\{\mathbf{C}^{(v)}\}_{v=1}^{V}, \mathbf{S}, \mathcal{E}, \mathcal{J}, \mathcal{G}, \mathcal{M}, \mathcal{N}) = \\ &\sum_{v=1}^{V} \mu^{(v)} \left\| \mathbf{S} - \mathbf{C}^{(v)} \right\|_{\mathrm{F}}^2 + \beta \|\mathcal{J}\|_{\mathrm{ETR}} + \gamma \|\mathcal{G}\|_{\mathrm{THLN}} \\ &+ 2\lambda \mathrm{Tr}(\mathbf{F}^\top \mathbf{L_S} \mathbf{F}) + \langle \mathcal{M}, \mathcal{C} - \mathcal{J} \rangle + \frac{\rho}{2} \|\mathcal{C} - \mathcal{J}\|_{\mathrm{F}}^2 \\ &+ \langle \mathcal{N}, \mathcal{A} - \mathcal{C} - \mathcal{G} \rangle + \frac{\varphi}{2} \|\mathcal{A} - \mathcal{C} - \mathcal{G}\|_{\mathrm{F}}^2, \end{aligned} \tag{6}$$

where the tensors $\mathcal{M}$ and $\mathcal{N}$ are Lagrange multipliers, and φ and ρ are penalty parameters to control convergence. The solution approach involves decomposing the overall optimization problem into five distinct subproblems.

S-subproblem: With all other variables fixed, the optimization problem for $\mathbf{S}$ is formulated as:

$$\begin{aligned} &\min_{\mathbf{S}} \sum_{v=1}^{V} \mu^{(v)} \left\| \mathbf{S} - \mathbf{C}^{(v)} \right\|_F^2 + 2\lambda \mathrm{Tr}(\mathbf{F}^\top \mathbf{L_S} \mathbf{F}) \\ &\text{s.t.} \quad \mathbf{s}_i^\top \mathbf{1} = 1, \mathbf{S} \geq 0. \end{aligned} \tag{7}$$

By defining a distance vector $\mathbf{a}_i \in \mathbb{R}^n$, where the j-th entry is $a_{ij} = \|\mathbf{f}_i - \mathbf{f}_j\|_2^2$, the problem can be decoupled into N independent subproblems. For the i-th row $\mathbf{s}_i$, the optimization is equivalent to finding the Euclidean projection of a vector onto the probability simplex. According to the KKT conditions [Boyd and Vandenberghe, 2004], the closed-form solution is derived as:

$$\mathbf{s}_i = \left(\left(\sum_{v=1}^{V} \mu^{(v)} \mathbf{c}_i^{(v)} - \frac{\lambda}{2} \mathbf{a}_i \right) \Big/ \sum_{v=1}^{V} \mu^{(v)} + \phi \right)_+, \tag{8}$$

where ϕ is the Lagrangian multiplier associated with the constraint $\mathbf{s}_i^\top \mathbf{1} = 1$.

F-subproblem: With all other variables fixed, the subproblem for $\mathbf{F}$ is formulated as follows:

$$\begin{aligned} &\min_{F} \mathrm{Tr}\left(\mathbf{F}^\top \mathbf{L_S} \mathbf{F}\right), \\ &\text{s.t.} \quad \mathbf{F} \in \mathbb{R}^{n \times k}, \mathbf{F}^\top \mathbf{F} = \mathbf{I}. \end{aligned} \tag{9}$$

This constitutes a classic spectral problem, where the corresponding solution can be derived by computing the k eigenvectors of $\mathbf{L_S}$ that correspond to its k smallest eigenvalues.

$\mathbf{C}^{(v)}$-subproblem: With all other variables fixed, the subproblem for $C^{(v)}$ is formulated as follows:

$$\min_{\mathbf{C}^{(v)}} \mu^{(v)} \left\| \mathbf{S} - \mathbf{C}^{(v)} \right\|_{\mathrm{F}}^2 + \frac{\rho}{2} \left\| \mathbf{C}^{(v)} - \mathbf{J}^{(v)} + \frac{\mathbf{M}^{(v)}}{\rho} \right\|_{\mathrm{F}}^2 + \frac{\varphi}{2} \left\| \mathbf{A}^{(v)} - \mathbf{C}^{(v)} - \mathbf{G}^{(v)} + \frac{\mathbf{N}^{(v)}}{\varphi} \right\|_{\mathrm{F}}^2 . \tag{10}$$

Since the objective is smooth and convex, the solution follows from the first-order optimality condition.

$$\mathbf{C}^{(v)} = \frac{2\mu^{(v)}\mathbf{S} + \rho\mathbf{J}^{(v)} - \mathbf{M}^{(v)} + \varphi(\mathbf{A}^{(v)} - \mathbf{G}^{(v)}) + \mathbf{N}^{(v)}}{2\mu^{(v)} + \rho + \varphi}. \tag{11}$$

$\mathcal{G}$-subproblem: With all other variables fixed, the subproblem for $\mathcal{G}$ is formulated as follows:

$$\min_{\mathcal{G}} \frac{\gamma}{\varphi} \|\mathcal{G}\|_{\mathrm{THLN}} + \frac{1}{2} \left\| \mathcal{G} - (\mathcal{A} - \mathcal{C} + \frac{\mathcal{N}}{\varphi}) \right\|_F^2 . \tag{12}$$

Theorem 1. *Suppose a third-order tensor $\mathcal{H} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with t-SVD $\mathcal{H} = \mathcal{U} * \mathcal{S} * \mathcal{V}^\top$, and $\alpha > 0$. The Tensor High-Rank Logarithmic Norm Minimization problem can be described as follows,*

$$\min_{\mathcal{G}} \alpha \|\mathcal{G}\|_{\mathrm{THLN}} + \frac{1}{2} \|\mathcal{G} - \mathcal{H}\|_F^2 . \tag{13}$$

The optimal solution $\mathcal{G}^*$ can then be derived as follows:

$$\mathcal{G}^* = \mathcal{U} * ifft(Prox_{f,\alpha}(\mathcal{S}_f), [], 3) * \mathcal{V}^T, \tag{14}$$

where $ifft(\mathrm{Prox}_{f,\beta}(\mathcal{S}_f), [], 3) \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is an f-diagonal tensor, and $\mathrm{Prox}_{f,\alpha}(\mathcal{S}_f^k(i,i))$ satisfies,

$$\min_{x \geq 0} \frac{1}{2}(x - \mathcal{S}_f^k(i,i))^2 + \alpha f(x), \tag{15}$$

where $f(x) = -x \log(1 + x)$.

Given that Eq. (15) in Theorem 1 is a combination of concave and convex functions, we can utilize DC Programming [Tao and An, 1997] to derive a closed-form solution:

$$\tau^{iter+1} = \left(S_f^k(i,i) - \frac{\gamma \cdot \partial f(\tau^{iter})}{\varphi} \right)_+, \tag{16}$$

where $\tau = Prox_{f,\alpha}(\mathcal{S}_f^k(i,i))$, and $iter$ is the iteration index.

$\mathcal{J}$-subproblem: With all other variables fixed, the subproblem for $\mathcal{J}$ is formulated as follows:

$$\min_{\mathcal{J}} \beta \|\mathcal{J}\|_{ETR} + \frac{\rho}{2} \left\| \mathcal{J} - (\mathcal{C} + \frac{\mathcal{M}}{\rho}) \right\|_F^2 \tag{17}$$

Similar to the $\mathcal{G}$-subproblem, the optimization of $\mathcal{J}$ can be solved in the Fourier domain following the same derivation strategy as Theorem 1. The solution is given by:

$$\tau^{iter+1} = \left(S_f^k(i,i) - \frac{\beta \cdot \partial g(\tau^{iter})}{\rho} \right)_+ . \tag{18}$$

where $g(x) = \frac{e^{\delta^2} x}{\delta + x}$, and $iter$ is the iteration index.

Finally, the Lagrange multipliers and penalty parameters are updated according to the following rules:

$$\begin{cases} \mathcal{M} = \mathcal{M} + \rho(\mathcal{C} - \mathcal{J}), \\ \mathcal{N} = \mathcal{N} + \varphi(\mathcal{E} - \mathcal{G}), \\ \rho = \eta_\rho \rho, \varphi = \eta_\varphi \varphi, \end{cases} \tag{19}$$

where $\eta_\rho, \eta_\varphi > 1$ are used to accelerate convergence.

Algorithm 1 Optimization Algorithm of HTS-CMGL

Input:
Multi-view data $\{\mathbf{X}^{(v)}\}_{v=1}^V$, cluster number c, trade-off parameters β, γ and δ, self-tuned parameter λ, the maximum number of iterations $t = 30$; Initialize the weight factor $\mu^{(i)} = \frac{1}{v}$ for each view,$iter = 1$,$S = \sum_{v=1}^V \mu^{(v)} A^{(v)}$;
Output:
Consensus matrix $\mathbf{S}$ and clustering labels $\mathbf{y}$.
1: Initialize $\mathbf{A}^{(v)}$, $\mathbf{C}^{(v)}$, $\mathbf{E}^{(v)}$, $\mathbf{S}$, $\mathbf{F}$, weights $\boldsymbol{\mu}$, and parameters ρ, φ;
2: **while** not converged **do**
3: Optimize $\mathbf{S}$ by Eq. (8);
4: Optimize $\mathbf{F}$ and rank parameter λ by Eq. (9);
5: Optimize $\mathbf{C}^{(v)}$ by Eq. (11);
6: Optimize $\mathcal{G}$ by Eq. (16);
7: Optimize $\mathcal{J}$ by Eq. (18);
8: Optimize $\mathcal{M}, \mathcal{N}$ and ρ, φ by Eq. (19);
9: **end while**
10: Perform spectral clustering on $\mathbf{S}$ to obtain final clustering results $\mathbf{y}$.

3.3 Computation Complexity Analysis

The computational complexity of HTS-CMGL is mainly determined by the optimization of the subproblems described in Section 3.2. Specifically, the time complexities for updating the variables $\mathbf{C}^{(v)}, \mathbf{S}, \mathbf{F}, \mathcal{J}, \mathcal{G}$ are $O(VN^2)$, $O(N^2)$, $O(N^3)$, $O(VN^3 + VN^2 \log V)$, and $O(VN^2)$, respectively, where N denotes the number of samples and V denotes the number of views. Since $N \gg V$ in most practical cases, the overall computational complexity of HTS-CMGL is therefore dominated by $O(tVN^3)$, where t denotes the number of iterations. In practice, the algorithm usually converges within about 30 iterations on all datasets.

4 Experiments

4.1 Experimental Setting

Dataset	Classes	Samples	Feature Dim.
100leaves	100	1600	64/64/64
NGs	5	500	2000/2000/2000
Caltech101-7	7	1474	435/798/52/34/35/64/56
Mfeat	10	2000	76/216/64/6/47/240
COIL20	20	1440	1024/3304/6750
MSRC-v6	7	210	1320/48/512/100/256/210
COIL100	100	7200	30/99/30

Table 1: The characteristics of our employed datasets

Data sets: To evaluate the effectiveness of HTS-CMGL, we conducted experiments on seven widely used multi-view clustering benchmark datasets: 100leaves [Mallah *et al.*, 2013], NGs [Lang, 1995], Caltech101-7 [Fei-Fei *et al.*, 2004], Mfeat [van Breukelen *et al.*, 1998], COIL20 [Nene *et al.*, 1996b], COIL100 [Nene *et al.*, 1996a], and MSRC-v6 [Winn and Jojic, 2005]. Detailed descriptions of these datasets are provided in Table 1.

Dataset	Metric(%)	GMC	CDMGC	FMICE	OMVFC-LICAG	IWTSN-FMGC	TAD-MVC	MCHBG	OURS
100leaves	ACC	82.38 ± 0.00	89.94 ± 0.88	83.39 ± 1.33	49.88± 0.00	87.92 ± 1.15	75.06 ± 0.00	87.94 ± 0.00	**96.00± 0.00**
	NMI	90.25 ± 0.00	95.30 ± 0.51	93.25 ± 0.60	71.17 ± 0.00	95.64 ± 0.41	87.24 ± 0.00	93.15 ± 0.00	**98.50 ± 0.00**
	PUR	85.06 ± 0.00	91.06 ± 0.64	92.03 ± 1.71	51.56 ± 0.00	90.20 ± 0.95	76.50 ± 0.00	89.13 ± 0.00	**96.88 ± 0.00**
	ARI	49.74 ± 0.00	84.13 ± 5.18	79.85 ± 1.48	31.51 ± 0.00	85.23 ± 1.37	65.20 ± 0.00	74.17 ± 0.00	**94.92 ± 0.00**
NGs	ACC	98.20 ± 0.00	77.08 ± 0.25	40.68 ± 5.40	59.00 ± 0.00	59.69 ± 0.15	98.40 ± 0.00	88.00 ± 0.00	**98.80 ± 0.00**
	NMI	93.92 ± 0.00	74.83 ± 0.58	25.69 ± 4.93	40.77 ± 0.00	34.37 ± 0.34	94.61 ± 0.00	72.89 ± 0.00	**95.82 ± 0.00**
	PUR	98.20 ± 0.00	77.48 ± 0.25	66.36 ± 4.10	59.00 ± 0.00	59.69 ± 0.15	98.40 ± 0.00	88.00 ± 0.00	**98.80 ± 0.00**
	ARI	95.54 ± 0.00	69.74 ± 0.75	14.72 ± 4.57	31.37 ± 0.00	29.50 ± 0.27	96.03 ± 0.00	71.81 ± 0.00	**97.01 ± 0.00**
Caltech101-7	ACC	69.20 ± 0.00	51.68±11.69	54.49 ± 1.09	44.64 ± 0.00	54.61 ± 0.00	54.82 ± 0.00	65.74 ± 0.00	**81.07 ± 0.00**
	NMI	**60.56 ± 0.00**	23.67±16.02	52.84 ± 1.70	35.43 ± 0.00	2.04 ± 0.03	1.91 ± 0.00	56.23 ± 0.00	54.01 ± 0.00
	PUR	88.47 ± 0.00	61.80±12.30	62.93 ± 1.66	82.29 ± 0.00	54.76 ± 0.02	55.16 ± 0.00	**89.82 ± 0.00**	85.75 ± 0.00
	ARI	**59.43 ± 0.00**	5.94±23.26	44.64 ± 1.75	31.08 ± 0.00	2.19 ± 0.00	0.63 ± 0.00	45.90 ± 0.00	59.40 ± 0.00
Mfeat	ACC	88.20 ± 0.00	84.43±0.04	84.97 ± 4.11	67.10 ± 0.00	67.60 ± 0.00	91.15 ± 0.00	83.25 ± 0.00	**96.55 ± 0.00**
	NMI	89.32 ± 0.00	89.62 ± 0.13	86.30 ± 2.04	64.78 ± 0.00	64.03 ± 0.00	83.93 ± 0.00	83.08 ± 0.00	**92.65 ± 0.00**
	PUR	88.20 ± 0.00	88.20 ± 0.03	90.23 ± 2.78	67.10 ± 0.00	69.25 ± 0.00	91.15 ± 0.00	84.05± 0.00	**96.55 ± 0.00**
	ARI	84.96 ± 0.00	83.16 ± 0.09	81.30 ± 3.15	54.01 ± 0.00	52.44 ± 0.00	81.58 ± 0.00	77.15± 0.00	**92.47 ± 0.00**
COIL20	ACC	79.10 ± 0.00	86.52 ± 0.75	78.49 ± 3.41	51.35 ± 0.00	76.03 ± 1.75	75.49 ± 0.00	87.64 ± 0.00	**100.00 ± 0.00**
	NMI	91.89 ± 0.00	94.41 ± 0.11	89.40 ± 1.64	69.57 ± 0.00	83.91 ± 0.99	82.46 ± 0.00	94.38 ± 0.00	**100.00 ± 0.00**
	PUR	84.79 ± 0.00	89.94 ± 0.07	88.93 ± 2.13	56.60 ± 0.00	76.98 ± 1.30	75.76 ± 0.00	89.93 ± 0.00	**100.00 ± 0.00**
	ARI	78.19 ± 0.00	83.81 ± 0.57	76.38 ± 4.07	47.08 ± 0.00	70.86 ± 1.86	68.99 ± 0.00	84.51 ± 0.00	**100.00 ± 0.00**
MSRC-v6	ACC	89.52 ± 0.00	87.57 ± 0.15	92.57 ± 0.54	70.00 ± 0.00	84.02 ± 0.63	91.43 ± 0.00	91.43 ± 0.00	**94.76± 0.00**
	NMI	81.64 ± 0.00	79.17 ± 0.00	85.46 ± 1.24	57.13 ± 0.00	73.25 ± 0.92	84.05 ± 0.00	83.16 ± 0.00	**89.91 ± 0.00**
	PUR	89.52 ± 0.00	87.57 ± 0.15	92.57 ± 0.54	70.48 ± 0.00	84.02 ± 0.63	91.43 ± 0.00	91.43 ± 0.00	**94.76 ± 0.00**
	ARI	76.68 ± 0.00	72.65 ± 0.22	83.23 ± 1.14	48.28 ± 0.00	67.93 ± 1.08	80.69 ± 0.00	80.86 ± 0.00	**87.95 ± 0.00**
COIL100	ACC	71.42 ± 0.00	73.26 ± 0.39	76.66 ± 2.07	78.43 ± 0.00	28.12 ± 0.35	31.17 ± 0.00	67.90 ± 0.00	**81.43± 0.00**
	NMI	86.53 ± 0.00	87.96 ± 0.63	92.05 ± 0.48	90.28 ± 0.00	48.97 ± 0.11	49.84 ± 0.00	79.7 ± 0.00	**93.29± 0.00**
	PUR	78.22 ± 0.00	79.31 ± 0.45	**91.13 ± 1.52**	80.90 ± 0.00	31.01 ± 0.26	33.00 ± 0.00	72.51 ± 0.00	84.39 ± 0.00
	ARI	32.15 ± 0.00	39.32 ± 2.64	**75.25 ± 1.26**	73.05 ± 0.00	7.85 ± 0.07	25.67 ± 0.00	22.93 ± 0.00	71.78 ± 0.00

Table 2: Results (mean±std) of our proposed method and other compared methods on seven data sets.

Baselines: We compare HTS-CMGL with several representative multi-view clustering methods, including GMC [Wang *et al.*, 2020], CDMGC [Huang *et al.*, 2022], FMICE [Huang *et al.*, 2023], OMVFC-LICAG [Zhang *et al.*, 2024], IWTSN-FMGC [Sun *et al.*, 2024], TAD-MVC [Chen *et al.*, 2025], and MCHBG [Zhao *et al.*, 2025]. All baseline methods are implemented using the parameter settings recommended in their original papers.

Evaluation Metrics: Before clustering, all datasets are normalized to ensure consistent feature scales across different views. Clustering performance is evaluated using Accuracy (ACC), Normalized Mutual Information (NMI), Purity (PUR), and Adjusted Rand Index (ARI). These metrics are widely used in multi-view clustering, and higher values generally indicate better performance.

Parameter Settings: The proposed HTS-CMGL involves four hyperparameters. To balance model flexibility and computational efficiency, we empirically fix $\beta = 10^{-1}$ and $\delta = 10^{-3}$, since the model is relatively insensitive to these two parameters. We therefore focus on tuning the remaining key parameters: the trade-off parameter γ and the Laplacian rank constraint parameter λ. Specifically, γ is selected from $\{10^{-3}, 10^{-2}, \ldots, 100\}$, while λ is tuned over $\{1, 2, \ldots, 10\}$. All methods are run 10 times, and we report the mean results with standard deviations. Experiments are conducted on a desktop computer with an Intel Core i9-14900K (3.20 GHz) CPU, 128 GB RAM, and MATLAB R2024a.

4.2 Experimental Results

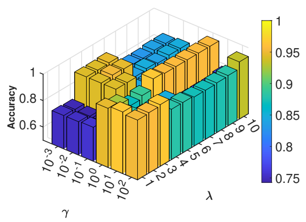
Table 2 presents a comprehensive quantitative evaluation of the proposed HTS-CMGL method against seven state-of-the-art baseline algorithms across seven diverse benchmark datasets. To provide a rigorous assessment of clustering quality, we report the mean values and standard deviations for four standard evaluation metrics. The best results are highlighted in **bold**, and the second-best results are underlined.

- The results show that HTS-CMGL consistently achieves strong performance across all datasets. Notably, it attains perfect clustering on COIL20 and maintains clear advantages on the remaining benchmarks. In terms of accuracy (ACC), HTS-CMGL ranks first among all compared methods, with the second-best results underlined in Table 2. It outperforms the closest baselines by margins of 6.06%, 0.40%, 11.87%, 5.40%, 12.36%, 2.19%, and 3.00% on the 100leaves, NGs, Caltech101-7, Mfeat, COIL20, MSRC-v6, and COIL100 datasets, respectively. These results demonstrate the effectiveness of HTS-CMGL. Moreover, its stable performance on diverse modalities, including text, object images, and video, indicates strong robustness and generalization ability across different scenarios.

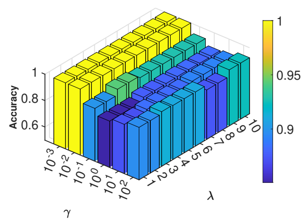
- Compared with tensor-based graph methods such as IWTSN-FMGC and MCHBG, HTS-CMGL achieves consistently better performance across all datasets. This advantage mainly stems from its explicit modeling of sample heterogeneity, together with the joint exploitation of high-order consistency and high-order specificity. In this way, by balancing these two aspects, the model effectively decouples consistency from diversity and yields a cleaner consensus graph.

5 Further Analysis

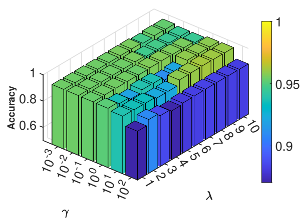
5.1 Parameter Sensitivity



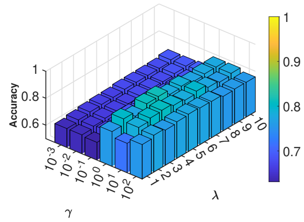
(a) MSRC-v6



(b) COIL20



(c) 100leaves



(d) COIL100

Figure 3: Parameters Analysis: The clustering performances (ACC) with different parameters γ and λ on four data sets.

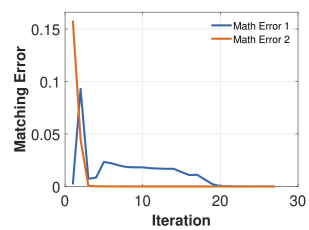
We perform a sensitivity analysis of HTS-CMGL with respect to two key hyperparameters: the Laplacian constraint parameter λ and the trade-off parameter γ. Owing to page limits, we only report the clustering performance (ACC) on four datasets, including MSRC-v6, COIL20, 100leaves, and COIL100, in Figure 3. The results show that both λ and γ have a clear influence on clustering performance. We can observe that when λ takes $\{3,\cdots,8\}$ and γ in $\{10^{-2},\cdots,10^{1}\}$, the ACC remains high. This phenomenon shows that the performance of our HTS-CMGL is stable across a wide range of parameters, and the stability of the parameters also directly illustrates the efficiency and robustness of HTS-CMGL.

5.2 Convergence Analysis

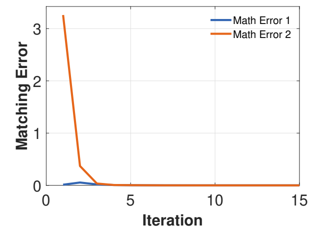
To examine the convergence behavior of HTS-CMGL, we conduct convergence experiments and present the results in this section. We adopt the matching errors of the consistency and diversity variables as convergence criteria. Specifically, the algorithm is considered converged when $\|\mathcal{J}-\mathcal{C}\|_\infty < 10^{-7}$ and $\|\mathcal{G}-\mathcal{E}\|_\infty < 10^{-7}$. For convenience, we denote $\|\mathcal{J}-\mathcal{C}\|_\infty$ as Math Error 1 and $\|\mathcal{G}-\mathcal{E}\|_\infty$ as Math Error 2. The variation curves of these errors on benchmark datasets, including COIL20, MSRC-v6, Mfeat and COIL100, are shown in Figure 4. HTS-CMGL converges rapidly on all datasets, with the errors decreasing sharply in the first few iterations and then remaining stable. The observed fast and stable convergence across different datasets indicates the excellent convergence of our HTS-CMGL.

5.3 Ablation Study

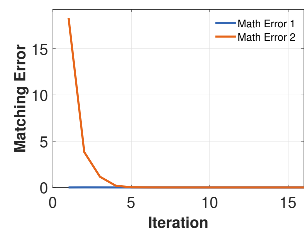
To investigate the contribution of ETR and THLN, we conducted an ablation study with two variants. In the first variant (Case 1), we replaced the ETR constraint with the conventional TNN to evaluate the effectiveness of the enhanced rank approximation. In the second variant (Case 2), we removed the THLN module to examine the necessity of modeling high-order specificity. Table 3 reports the clustering accuracy under these settings on four benchmark datasets. The complete model consistently outperforms both variants. Removing either component leads to performance degradation, which is evident on complex datasets, indicating that these constraints are crucial for maintaining stability and improving accuracy.



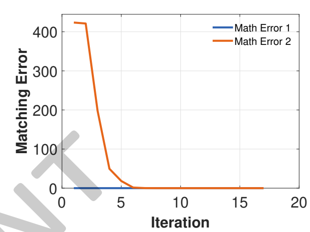
(a) COIL20



(b) MSRC-v6



(c) Mfeat



(d) COIL100

Figure 4: Convergence Analysis: The stop criteria (i.e., ME1 and ME2) variation curves on four data sets.

case1	case2	MSRC-v6	COIL20	100leaves	COIL100
×	✓	93.81	98.26	95.13	80.47
✓	×	94.28	95.14	94.25	66.81
✓	✓	**94.76**	**100.00**	**96.00**	**81.43**

Table 3: ACC(%) of ablation experiments

6 Conclusion

In this article, we propose High-Rank Tensor Specificity Induced Cooperative Multi-View Graph Learning (HTS-CMGL) to address the limitations of existing multi-view clustering methods in modeling heterogeneous sample diversity and high-order specificity. Specifically, HTS-CMGL tackles these issues through a novel learning framework. The core innovation lies in the synergistic integration of ETR and THLN, which allows the model to simultaneously enforce low-rank consistency and preserve fine-grained, high-rank diversity patterns. This separation mechanism effectively prevents locally co-occurring, view-specific structures from leaking into the global consensus graph. Extensive results demonstrate that our proposed HTS-CMGL is significantly superior to other state-of-the-art methods.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 62306020), the Young Elite Scientist Sponsorship Program by BAST (No. BYESS2024199), Beijing Natural Science Foundation (No. L244009), the Central Guidance for Local Scientific and Technological Development Fund (No.2024ZY0124), and the National Key Research and Development Program of China (No. 2023YFB3107100).

References

[Bao *et al.*, 2025] Yichen Bao, Yuxuan Liu, Yu Duan, Jing Li, and Quanxue Gao. Multi-view clustering based on probabilistic tensor regression. In *Proceedings of the 33rd ACM International Conference on Multimedia*, pages 2169–2177, 2025.

[Boyd and Vandenberghe, 2004] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[Chen *et al.*, 2025] Yawei Chen, Huibing Wang, Mingze Yao, Jinjia Peng, Guangqi Jiang, and Jiqing Zhang. Scalable multi-view clustering based on tight anchor distribution. In *Proceedings of the 33rd ACM International Conference on Multimedia*, page 7739–7747, 2025.

[Fan, 1949] Ky Fan. On a theorem of weyl concerning eigenvalues of linear transformations i. *Proceedings of the National Academy of Sciences*, 35(11):652–655, 1949.

[Fei-Fei *et al.*, 2004] L Fei-Fei, Rob Fergus, and Pietro Perona. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. In *2004 Conference on Computer Vision and Pattern Recognition Workshop*, page 178, 2004.

[Feng *et al.*, 2025] Wei Feng, Dongyuan Wei, Qianqian Wang, and Bo Dong. Tensorial multi-view clustering with deep anchor graph projection. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pages 5127–5135, 2025.

[Guo *et al.*, 2023] Jipeng Guo, Yanfeng Sun, Junbin Gao, Yongli Hu, and Baocai Yin. Logarithmic schatten-pp norm minimization for tensorial multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(3):3396–3410, 2023.

[Guo *et al.*, 2024] Wei Guo, Hangjun Che, and Man-Fai Leung. Tensor-based adaptive consensus graph learning for multi-view clustering. *IEEE Transactions on Consumer Electronics*, 70(2):4767–4784, 2024.

[Huang *et al.*, 2022] Shudong Huang, Ivor W Tsang, Zenglin Xu, and Jiancheng Lv. Measuring diversity in graph learning: a unified framework for structured multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 34(12):5869–5883, 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*, 35(11):11388–11402, 2023.

[Huang *et al.*, 2024] Shudong Huang, Ivor W. Tsang, Zenglin Xu, and Jiancheng Lv. Cgdd: Multiview graph clustering via cross-graph diversity detection. *IEEE Transactions on Neural Networks and Learning Systems*, 35(3):4206–4219, 2024.

[Ji and Feng, 2025] Jintian Ji and Songhe Feng. Anchors crash tensor: efficient and scalable tensorial multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(4):2660–2675, 2025.

[Lang, 1995] Ken Lang. Newsweeder: learning to filter netnews. In *Proceedings of the 12th International Conference on Machine Learning*, pages 331–339, 1995.

[Li *et al.*, 2026] Zhuo Li, Gengyu Lyu, Yuena Lin, Ziang Chen, Zhiyuan Ma, Zhen Yang, and Zun Li. Ccahcl: Multi-level hypergraph contrastive learning for connected component awareness. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 23355–23363, 2026.

[Liang *et al.*, 2019] Youwei Liang, Dong Huang, and Chang-Dong Wang. Consistency meets inconsistency: a unified graph learning framework for multi-view clustering. In *2019 IEEE International Conference on Data Mining*, pages 1204–1209, 2019.

[Lin *et al.*, 2011] Zhouchen Lin, Risheng Liu, and Zhixun Su. Linearized alternating direction method with adaptive penalty for low-rank representation. In *Proceedings of the 25th International Conference on Neural Information Processing Systems*, 2011.

[Lin *et al.*, 2025] Yuena Lin, Haichun Cai, Jun-Yi Hang, Haobo Wang, Zhen Yang, and Gengyu Lyu. Mitigating local cohesion and global sparseness in graph contrastive learning with fuzzy boundaries. In *Proceedings of the 42nd International Conference on Machine Learning*, pages 37745–37761, 2025.

[Lin *et al.*, 2026] Yuena Lin, Hao Wei, Hai-Chun Cai, Bohang Sun, Tao Yang, Zhen Yang, and Gengyu Lyu. Caligcl: Calibrated graph contrastive learning via partitioned similarity and consistency discrimination. *Advances in Neural Information Processing Systems*, pages 126748–126774, 2026.

[Lyu *et al.*, 2024] Gengyu Lyu, Weiqi Kang, Haobo Wang, Zheng Li, Zhen Yang, and Songhe Feng. Common-individual semantic fusion for multi-view multi-label learning. In *Proceedings of the 33rd International Joint Conference on Artificial Intelligence*, pages 4715–4723, 2024.

[Mallah *et al.*, 2013] Charles Mallah, James Cope, and James Orwell. Plant leaf classification using probabilistic integration of shape, texture and margin features. In *Proceedings of the IASTED International Conference on Signal Processing, Pattern Recognition and Applications*, pages 279–286, 2013.

[Nene *et al.*, 1996a] Sameer A Nene, Shree K Nayar, and Hiroshi Murase. Columbia object image library (coil-

100). Technical Report CUCS-006-96, Columbia University, 1996.

[Nene *et al.*, 1996b] Sameer A Nene, Shree K Nayar, and Hiroshi Murase. Columbia object image library (coil-20). Technical Report CUCS-005-96, Columbia University, 1996.

[Peng *et al.*, 2025] Xin Peng, Bowen Liu, Renxiang Guan, and Wenxuan Tu. Multi-view graph clustering with dual structure awareness for remote sensing data. In *Proceedings of the 33rd ACM International Conference on Multimedia*, pages 2313–2322, 2025.

[Pfister *et al.*, 2018] Niklas Pfister, Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. Kernel-based tests for joint independence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(1):5–31, 2018.

[Sun *et al.*, 2024] Yinghui Sun, Xingfeng Li, Quansen Sun, Min-Ling Zhang, and Zhenwen Ren. Improved weighted tensor schatten p-norm for fast multi-view graph clustering. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pages 1427–1436, 2024.

[Sun *et al.*, 2025a] Bohang Sun, Yongjian Deng, Yuena Lin, Qiuru Hai, Zhen Yang, and Gengyu Lyu. Graph consistency and diversity measurement for federated multi-view clustering. In *Proceedings of the AAAI conference on Artificial Intelligence*, pages 20663–20671, 2025.

[Sun *et al.*, 2025b] Bohang Sun, Yuena Lin, Tao Yang, Zhen Zhu, Zhen Yang, and Gengyu Lyu. Af-umc: an alignment-free fusion framework for unaligned multi-view clustering. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, pages 104157–104179, 2025.

[Sun *et al.*, 2025c] Weijun Sun, Chaoye Li, Kangsheng Wu, et al. Multi-view graph clustering with dually enhanced tensor rank minimization and diverse separation of inconsistent information. *Neural Networks*, 189:107550, 2025.

[Tang *et al.*, 2025] Chuan Tang, Miaomiao Li, Jun Wang, Chang Tang, Jiahe Jiang, Tianyi Wang, En Zhu, and Xinwang Liu. Multi-view clustering via high-order bipartite graph learning and tensor low-rank representation. *IEEE Transactions on Knowledge and Data Engineering*, 37(11):6521–6534, 2025.

[Tao and An, 1997] Pham Dinh Tao and Le Thi Hoai An. Convex analysis approach to dc programming: theory, algorithms and applications. *Acta Mathematica Vietnamica*, 22(1):289–356, 1997.

[van Breukelen *et al.*, 1998] Marcel van Breukelen, Robert P. W. Duin, David M. J. Tax, and J. E. Den Hartog. Handwritten digit recognition by combined classifiers. *Kybernetika*, 34(4):381–386, 1998.

[Wang *et al.*, 2020] Hao Wang, Yan Yang, and Bing Liu. Gmc: Graph-based multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 32(6):1116–1129, 2020.

[Wang *et al.*, 2025] Qianqian Wang, Zihao Zhang, Wei Feng, Zhiqiang Tao, and Quanxue Gao. Contrastive multi-view subspace clustering via tensor transformers autoencoder. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 21207–21215, 2025.

[Winn and Jojic, 2005] John Winn and Nebojsa Jojic. Locus: learning object classes with unsupervised segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 756–763, 2005.

[Xia *et al.*, 2023] Wei Xia, Quanxue Gao, Qianqian Wang, Xinbo Gao, Chris Ding, and Dacheng Tao. Tensorized bipartite graph learning for multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4):5187–5202, 2023.

[Xie *et al.*, 2018] Yuan Xie, Dacheng Tao, Wensheng Zhang, Yan Liu, Lei Zhang, and Yanyun Qu. On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *International Journal of Computer Vision*, 126(11):1157–1179, 2018.

[Yang *et al.*, 2025] Jie Yang, Wei Chen, Feng Liu, Peng Zhou, Zhongli Wang, Xinyan Liang, and Bingbing Jiang. Multi-view clustering via multi-granularity ensemble. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pages 6794–6802, 2025.

[Yang *et al.*, 2026] Jie Yang, Cheng-You Lu, Zhongli Wang, Hsiang-Ting Chen, Guang-Kui Xu, Chenglong Zhang, Shuting Dong, Xinyan Liang, and Bingbing Jiang. Multi-view clustering with granularity-aware pseudo supervision. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 27538–27546, 2026.

[Zhan *et al.*, 2019] Kun Zhan, Feiping Nie, Jing Wang, and Yi Yang. Multiview consensus graph clustering. *IEEE Transactions on Image Processing*, 28(3):1261–1270, 2019.

[Zhang *et al.*, 2024] Chuanbin Zhang, Long Chen, Zhaoyin Shi, and Weiping Ding. Latent information-guided one-step multi-view fuzzy clustering based on cross-view anchor graph. *Information Fusion*, 102:102025, 2024.

[Zhao *et al.*, 2025] Zihua Zhao, Ting Wang, Haonan Xin, Rong Wang, and Feiping Nie. Multi-view clustering via high-order bipartite graph fusion. *Information Fusion*, 113:102630, 2025.

[Zhen *et al.*, 2025] Long Zhen, Yipeng Liu, Yazhou Ren, and Ce Zhu. Slr-mvtc: smooth low-rank multi-view tensor clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 19095–19103, 2025.

[Zhong *et al.*, 2025] Qiyu Zhong, Yi Shan, Haobo Wang, Zhen Yang, and Gengyu Lyu. Tensorized multi-view multi-label classification via laplace tensor rank. In *Proceedings of the 42nd International Conference on Machine Learning*, pages 78522–78536, 2025.

[Zou *et al.*, 2025] Jiale Zou, Yan Chen, Bingbing Jiang, Peng Zhou, Liang Du, Lei Duan, and Yuhua Qian. Robust tensor learning with graph diffusion for scalable multi-view graph clustering. In *Proceedings of the 33rd ACM International Conference on Multimedia*, pages 2207–2215, 2025.