Consensus-Guided Incomplete Multi-view Clustering via Cross-view Affinities Learning

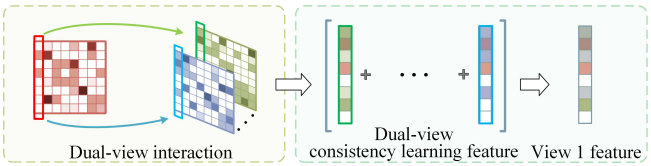
Qian Liu[1], Huibing Wang[1]*, Jinjia Peng[2], Yawei Chen[1], Mingze Yao[1], Xianping Fu[1], Yang Wang[3]
[1]School of Information Science and Technology, Dalian Martime University, Dalian, China
[2]School of Cyber Security and Computer, Hebei University, Baoding, China
[3]School of Computer Science and Information Engineering, Hefei University of Technology, China
{lq0307,huibing.wang,cyw,ymz0284,fxp}@dlmu.edu.cn, pengjinjia@hbu.edu.cn,
yangwang@hfut.edu.cn

Abstract

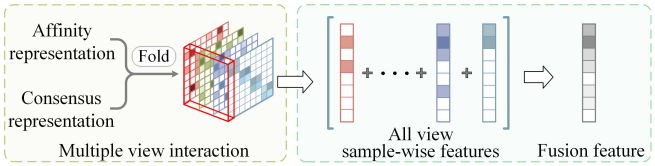
Incomplete multi-view clustering (IMC) has garnered substantial attention due to its capacity to handle unlabeled data. Existing methods predominantly explore pairwise consistency between every two views. However, such consistency is highly susceptible to missing samples and outliers within a certain view and thus deviates from the true clustering distribution. Moreover, dual-view interaction neglects the collaboration effects of multiple views, making it challenging to capture the holistic characteristics across views. In response to these issues, we propose a novel Consensus-Guided Incomplete Multi-view Clustering via Cross-view Affinities Learning (CAL). Specifically, CAL reconstructs views with available instances to mine sample-wise affinities and harness comprehensive content information within views. Subsequently, to extract clean structural information, CAL imposes a structured sparse constraint on the representation tensor to eliminate biased errors. Furthermore, by integrating the consensus representation into a representation tensor, CAL can employ high-order interaction of multiple views to depict the semantic correlation between views while acquiring a unified structural graph across multiple views. Extensive experiments on seven benchmark datasets demonstrate that CAL outperforms some state-of-the-art methods in clustering performance. The code is available at https://github.com/whbdmu/CAL.

1 Introduction

In the wake of the ongoing progression of information technology, the same object can be observed from multiple sources and engender multi-view data [Wang, 2021], [Fang *et al.*, 2023], [Wang *et al.*, 2023]. The consistent and complementary information embedded within such data is customarily harnessed to tackle the multi-view clustering (MVC) [Su *et al.*, 2024], [Tan *et al.*, 2024], [Long *et al.*, 2024a] task. However, in practical applications, it is often intractable to ensure the completeness of multi-view data [Wu *et al.*, 2024].



(a) Dual-view interaction for consistency learning



(b) Multiple view interaction for consensus learning

Figure 1: Comparison between traditional strategy and our strategy. (a) The traditional strategy focuses on pairwise interaction between the current view and others. (b) Our strategy fully considers the collaborative effects among multiple views and leverages consensus representation to regulate the learning of high-order consistency.

As a result, incomplete multi-view clustering (IMC) [Wang *et al.*, 2024a], [Long *et al.*, 2024b], [Yao *et al.*, 2025] has attracted considerable attention for its ability to partition missing data into their corresponding classes.

Owing to the remarkable success of multi-view subspace clustering (MVSC) as described in [Cao *et al.*, 2015], [Kang *et al.*, 2020], [Huang *et al.*, 2022], most IMC approaches [Li *et al.*, 2022], [Gu *et al.*, 2024] follow a pipeline similar to MVSC with a two-step strategy: constructing a unified affinity graph followed by applying spectral clustering, in which the construction of the affinity graph plays a pivotal role in the IMC process.

According to the way of affinity graph construction, existing IMC methods can be categorized into subspace clustering-based methods [Liang *et al.*, 2024], [Liu *et al.*, 2024a] and graph learning-based methods [Wang *et al.*, 2024b], [Chen *et al.*, 2025]. For instance, in [Wen *et al.*, 2020], missing view inference and the semantic consistency constraint were incorporated into MVSC to learn complete pairwise consistency between views. [Zhao *et al.*, 2023] applies the between-view consistency constraint to establish a joint graph completion framework and achieve pairwise consistency learning. The mentioned IMC methods prioritize

*Corresponding Author.

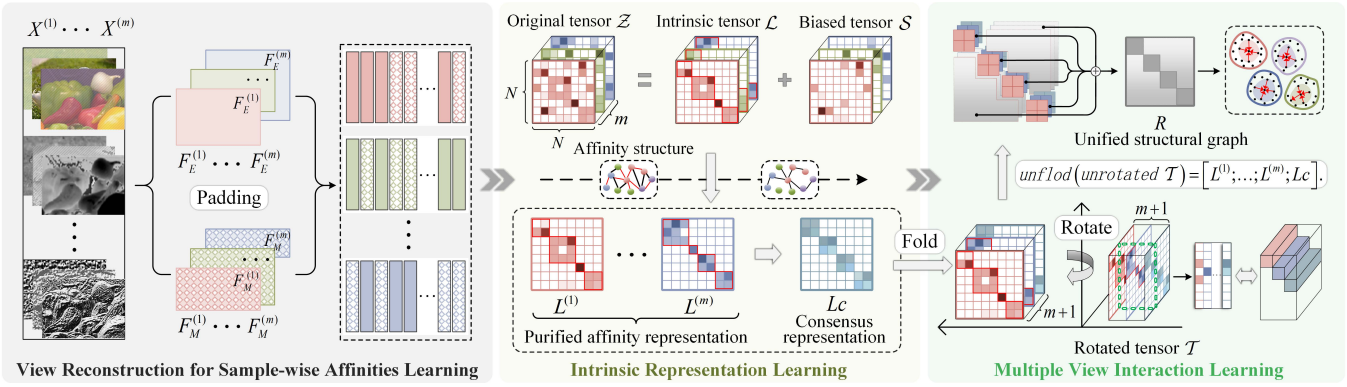


Figure 2: The framework of our proposed CAL, which learns a unified structural graph through the high-order interaction between content and structural information across multiple views, under the guidance of a consensus representation.

the recovery of missing instances to enhance the accuracy of affinity graphs, a principle that our method also follows. While these methods have achieved outstanding performance in the IMC task, they often rely on simple dual-view interaction to explore pairwise consistency, as depicted in Figure 1(a). This reliance may restrict the comprehensive exploration of holistic characteristics and clustering structure information across multiple views. Meanwhile, the pairwise consistency calculation between views tends to diverge from the true data clustering distribution affected by various noises.

To surmount the above weakness, this paper puts forward a novel cross-view affinities learning (CAL) framework for IMC, which learns a unified structural graph through the high-order interaction between content and structural information across multiple views, under the guidance of a consensus representation. Specifically, two padding matrices are introduced to extract comprehensive content information for each view and capture sample-wise affinities. Following this, CAL applies the structured sparsity constraint on acquired affinity representations to filter out erroneous connections between clusters arising from reconstruction errors and inherent noise. Furthermore, CAL decomposes view-specific affinity representations into a consensus representation and diverse representations with linear relationships. Unlike prior methods that rely on dual-view interaction, CAL stacks the consensus representation together with affinity representations into a tensor, making full use of the collaborative effects across multiple views to excavate and learn a structural graph with high-order semantic consistency. Eventually, to preserve the distinctiveness of each view, the view-specific representations and the learned structural graph jointly contribute to the computation to yield a unified structural graph. In summary, the flowchart of CAL is illustrated in Figure 2, and its main contributions can be encapsulated as follows:

- We propose a novel cross-view affinity learning method, which stacks the consensus representation and the affinity representations into a high-order tensor. It can efficiently learn a unified structural graph by considering collaborative effects across multiple views.
- By recovering missing instances and applying a structured sparse constraint to the representation tensor, CAL enhances the learning of comprehensive and precise content and structural information in each view, which facilitates the acquisition of a unified structural graph.
- To substantiate the superiority of high-order interaction over dual-view interaction, extensive experiments were conducted on seven benchmark datasets, and the comparison results demonstrated that our approach can achieve state-of-the-art performances in most cases.

2 The Proposed Method

In this section, we will delve into the details of the proposed CAL. As shown in Figure. 2, CAL aims to learn a unified structural graph based on the complete content information and clean structural information within each view via high-order interaction across multiple views.

Suppose $\{\mathbf{X}^{(v)}\}_{v=1}^{m}$ denote incomplete multi-view data with m views and can be divided into an existing instance set $\mathbf{X}_E^{(v)} \in \mathbb{R}^{d_v \times N_E^{(v)}}$ and missing instance set $\mathbf{X}_M^{(v)} \in \mathbb{R}^{d_v \times N_M^{(v)}}$, i.e., $\mathbf{X}^{(v)} = [\mathbf{X}_E^{(v)}, \mathbf{X}_M^{(v)}]$, where the feature dimension is denoted by d_v, and N represents the total number of instances, which is equivalent to the sum of $N_E^{(v)}$ and $N_M^{(v)}$.

Sample-wise Affinities Learning: In order to exploit the complete semantic information of each view to enhance the expressive ability of affinity representations, CAL introduces two padding matrices $\mathbf{F}_E^{(v)} \in \mathbb{R}^{N_E^{(v)} \times N}$ and $\mathbf{F}_M^{(v)} \in \mathbb{R}^{N_M^{(v)} \times N}$ that map existing and missing instances to respective positions in the complete feature matrix to reconstruct incomplete views, such that the v-th feature matrix can be rewritten as:

$$\mathbf{X}^{(v)} = \mathbf{X}_E^{(v)}\mathbf{F}_E^{(v)} + \mathbf{X}_M^{(v)}\mathbf{F}_M^{(v)} \tag{1}$$

where $\mathbf{F}_E^{(v)}$ is constructed by iteratively substituting the columns of the zero matrix $\mathbf{O}_E^{(v)} \in \mathbb{R}^{N_E^{(v)} \times N}$, which correspond to the indices of existing data in view v with the columns of the identity matrix $\mathbf{I}_E^{(v)} \in \mathbb{R}^{N_E^{(v)} \times N_E^{(v)}}$. Similarly, $\mathbf{F}_M^{(v)}$ is constructed by substituting the columns of the zero matrix $\mathbf{O}_M^{(v)} \in \mathbb{R}^{N_M^{(v)} \times N}$ corresponding to the indices of missing data in view v with the columns of the identity matrix $\mathbf{I}_M^{(v)} \in \mathbb{R}^{N_M^{(v)} \times N_M^{(v)}}$.

Then, adhering to the principle that each data point is amenable to linear representation by others, we carry out view reconstruction involving missing instances and formulate the objective function for sample-wise affinities learning as

$$\begin{aligned}\min_{\mathbf{X}_M^{(v)},\mathbf{Z}^{(v)}} &\sum_{v=1}^{m}\left(\left\|\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{Z}^{(v)}-\mathbf{B}_1^{(v)}\right\|_F^2+\left\|\mathbf{B}_1^{(v)}\right\|_1\right)\\ &s.t.\ \mathbf{X}^{(v)}=\mathbf{X}_E^{(v)}\mathbf{F}_E^{(v)}+\mathbf{X}_M^{(v)}\mathbf{F}_M^{(v)},\\ &\mathbf{Z}^{(v)}\geq 0,\mathbf{Z}^{(v)}\mathbf{1}=1,\mathrm{diag}(\mathbf{Z}^{(v)})=0.\end{aligned} \tag{2}$$

where $\mathbf{Z}^{(v)} \in \mathbb{R}^{N\times N}$ and $\mathbf{B}_1^{(v)} \in \mathbb{R}^{d_{(v)}\times N}$ are an affinity representation and error matrix for v-th view, respectively. In the above ways, CAL can effectively impute the missing data to mitigate the deviation caused by data imperfections. Building upon this, CAL further captures the richer content information for each view and explores the sample-wise affinities.

Intrinsic Representation learning: We regard the harmful information, induced by reconstruction error, outliers, and inherent noise concealed within views, as biased error. The affinity structure in $\mathbf{Z}^{(v)}$ is vulnerable to being compromised by such biased error, giving rise to erroneous inter-cluster connections. Directly utilizing these damaged affinity representations to learn consistency between views could lead to a divergence from the realistic clustering distribution of the data. To this end, we stack $\{\mathbf{Z}^{(v)}\}_{v=1}^m$ into a tensor $\mathcal{Z}$ and subsequently separate the biased error by decomposing $\mathcal{Z}$ as follows:

$$\mathcal{Z}=\mathcal{L}+\mathcal{S} \tag{3}$$

where the intrinsic tensor $\mathcal{L}$ encodes the real relationships among all samples, including available and absent instances in each view. The biased tensor $\mathcal{S}$ is utilized to describe the harmful information and the $\ell_{2,1}$-norm is employed to capture the biased error among views as

$$\begin{aligned}\min_{\substack{\mathcal{Z},\mathcal{S}\\ \mathbf{B}_2^{(v)},\mathbf{L}^{(v)},\mathbf{L}_c}} &\sum_{v=1}^{m}\lambda_1\left\|\mathbf{B}_2^{(v)}\right\|_1+\lambda_2\|\mathcal{S}\|_{2,1}\\ s.t.\ &\mathcal{Z}=\Phi(\mathbf{Z}^{(1)},\mathbf{Z}^{(2)},...,\mathbf{Z}^{(v)}),\ \mathcal{Z}=\mathcal{L}+\mathcal{S},\\ &\mathbf{L}_c=\mathcal{J}\left(\left\{\mathbf{L}^{(v)}\right\}_{v=1}^m,\left\{\mathbf{B}_2^{(v)}\right\}_{v=1}^m\right)\end{aligned} \tag{4}$$

where Φ is an operator used to construct a third-order tensor by aggregating all affinity representations. Moreover, the purified view-specific affinity representations $\{\mathbf{L}^{(v)}\}_{v=1}^m$ are decomposed into linearly related consensus representation $\mathbf{L}_c$ and diverse representation $\mathbf{B}_2^{(v)}$.

Multiple View Interaction Learning: Upon the successful acquisition of comprehensive content information and clean structural information, we prefer to conduct high-order interaction among multiple views under the guidance of consensus representation rather than dual-view interaction. The dual-view interaction typically concentrates solely on the associations between two views at a time and there are significant limitations in mining the holistic characteristics of multi-view data. In contrast, high-order interaction can fully leverage the collaborative effects of multiple views and effectively break through the information barriers formed by dual-view interaction. Based on the above description, we can arrive at the overall objective function of our method CAL as follows:

$$\begin{aligned}\min_{\substack{\mathbf{X}_M^v,\mathbf{Z}^{(v)},\mathbf{B}^{(v)},\\ \mathbf{L}^{(v)},\mathbf{L}_c,\mathcal{S}}} &\sum_{v=1}^{m}\left\|\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{Z}^{(v)}-\mathbf{B}_1^{(v)}\right\|_F^2\\ &+\sum_{v=1}^{m}\lambda_1\left\|\mathbf{B}^{(v)}\right\|_1+\lambda_2\|\mathcal{S}\|_{2,1}+\|\mathcal{T}\|_{\omega,\mathfrak{Sp}}^p\\ \text{s.t. } \mathbf{X}^{(v)}=&\mathbf{X}_E^{(v)}\mathbf{F}_E^{(v)}+\mathbf{X}_M^{(v)}\mathbf{F}_M^{(v)},\ \mathbf{Z}^{(v)}\geq 0,\ \mathbf{Z}^{(v)}\mathbf{1}=1,\\ &\mathrm{diag}(\mathbf{Z}^{(v)})=0,\ \mathcal{Z}=\Phi(\mathbf{Z}^{(1)},\mathbf{Z}^{(2)},...,\mathbf{Z}^{(v)}),\\ &\mathcal{Z}=\mathcal{L}+\mathcal{S},\ L_c=\mathcal{J}\left(\{\mathbf{L}^{(v)}\}_{v=1}^m,\{\mathbf{B}_2^{(v)}\}_{v=1}^m\right),\\ &\mathbf{B}^{(v)}=[\mathbf{B}_1^{(v)};\mathbf{B}_2^{(v)}],\ \mathcal{T}=\Phi\left(\mathbf{L}^{(v)m}_{v=1},\mathbf{L}_c\right)\end{aligned} \tag{5}$$

where λ_1 and λ_2 are the trade-off parameters of the corresponding terms in the model. The tensor $\mathcal{T}\in\mathbb{R}^{N\times(m+1)\times N}$ is a rotated tensor stacked with the affinity representations $\mathbf{L}^{(v)}$ and the consensus representation $\mathbf{L}_c$. Within this tensor, the consensus representation acts as a coefficient of consistency and is enforced by the weighted tensor Schatten p-norm constraint, it effectively promotes the exploration and reinforcement of consistent information across all views and we can acquire a consistent structural graph $\mathbf{L}_c$. After learning the optimal $\mathbf{L}_c$ and $\mathbf{L}^{(v)}$, the unified structural graph $\mathbf{R}$ considering consistency and specificity can be calculated by

$$\mathbf{R}=\frac{1}{m}\sum_{v=1}^{m}\frac{\left|\mathbf{L}^{(v)}\right|+\left|\mathbf{L}^{(v)T}\right|}{2}+\frac{\left|\mathbf{L}_c\right|+\left|\mathbf{L}_c^T\right|}{2}, \tag{6}$$

and then the final clustering result is gained by performing spectral clustering on $\mathbf{R}$.

3 Optimization

3.1 Optimization Algorithm

By introducing an auxiliary tensor variable $\mathcal{A}$, the augmented Lagrangian function of problem (5) can be formulated as:

$$\begin{aligned}L(\varphi)=&\sum_{v=1}^{m}\left\|\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{Z}^{(v)}-\mathbf{B}_1^{(v)}\right\|_F^2+\lambda_1\|\mathbf{B}^{(v)}\|_1\\ &+\lambda_2\|\mathcal{S}\|_{2,1}+\frac{\rho}{2}\|\mathcal{Z}-\mathcal{L}-\mathcal{S}+\frac{\mathcal{C}_1}{\rho}\|_F^2+\|\mathcal{A}\|_{\omega,\mathfrak{Sp}}^p\\ &\sum_{v=1}^{m}\frac{\rho}{2}\|\mathbf{L}^{(v)}-\mathbf{L}_c-\mathbf{B}_2^{(v)}+\frac{\mathbf{C}_2^{(v)}}{\rho}\|_F^2\end{aligned} \tag{7}$$

where $\varphi=\{\{\mathbf{X}_M^{(v)},\mathbf{Z}^{(v)},\mathbf{B}^{(v)},\mathbf{L}^{(v)}\}_{v=1}^m,\mathbf{L}_c,\mathcal{S},\mathcal{A}\}$ represents the set of variables to be optimized, $\rho>0$ is a penalty parameter, $\mathcal{Y}$, $\mathcal{C}_1$ and $\mathbf{C}_2^{(v)}$ are Lagrange multipliers. We will iteratively optimize each variable with the following details.

(1) $\mathbf{X}_M^{(v)}$ Step: Fixing other variables and considering the constraint $\mathbf{X}^{(v)}=\mathbf{X}_E^{(v)}\mathbf{F}_E^{(v)}+\mathbf{X}_M^{(v)}\mathbf{F}_M^{(v)}$, we can derive the optimal solution by setting the partial derivatives of the minimization problem to zero:

$$\mathbf{X}_M^{(v)}=\left(-\mathbf{X}_E^{(v)}\mathbf{F}_E^{(v)}\mathbf{P}^{(v)}\mathbf{F}_M^{(v)T}\right)\left(\mathbf{F}_M^{(v)}\mathbf{P}^{(v)}\mathbf{F}_M^{(v)T}\right)^{-1} \tag{8}$$

where $\mathbf{P}^{(v)} = (\mathbf{I} - \mathbf{Z}^{(v)})(\mathbf{I} - \mathbf{Z}^{(v)})^T$.

(2) $\mathbf{Z}^{(v)}$ Step: When all the other variables are fixed, $\mathbf{Z}^{(v)}$ can be obtained by solving the following sub-problem:

$$\begin{aligned} L(\mathbf{Z}^{(v)}) &= \left\|\mathbf{X}^{(v)} - \mathbf{X}^{(v)}\mathbf{Z}^{(v)} - \mathbf{B}_1^{(v)}\right\|_F^2 \\ &+ \frac{\rho}{2}\left\|\mathbf{Z}^{(v)} - \mathbf{L}^{(v)} - \mathbf{S}^{(v)} + \frac{\mathbf{C}_1^v}{\rho}\right\|_F^2 \\ s.t.\ \mathbf{Z}^{(v)} &\geq 0,\ \mathbf{Z}^{(v)}\mathbf{1} = 1,\ \mathrm{diag}(\mathbf{Z}^{(v)}) = 0 \end{aligned} \tag{9}$$

By setting the partial derivative $\partial L\left(\mathbf{Z}^{(v)}\right) \Big/ \partial \mathbf{Z}^{(v)} = 0$, we can obtain that:

$$\begin{aligned} \widetilde{\mathbf{Z}}^{(v)} &= \left(\mathbf{X}^{(v)T}\mathbf{X}^{(v)} + \rho\mathbf{I}\right)^{-1} \\ &\left(\mathbf{X}^{(v)T}\mathbf{X}^{(v)} - \mathbf{X}^{(v)T}\mathbf{B}_1^{(v)} + \rho\mathbf{L}^{(v)} + \rho\mathbf{S}^{(v)} - \mathbf{C}_1^{(v)}\right) \end{aligned} \tag{10}$$

Then, the optional $\mathbf{Z}^{(v)}$ is obtained by solving

$$\mathbf{Z}^{(v)} = \underset{\substack{\mathbf{Z}^{(v)}\geq 0,\, \mathbf{Z}^{(v)}\mathbf{1}=1,\\ \mathrm{diag}\left(\mathbf{Z}^{(v)}\right)=0}}{\arg\min} \left\|\mathbf{Z}^{(v)} - \widetilde{\mathbf{Z}}^{(v)}\right\|_F^2 \tag{11}$$

The optional solution is [Nie *et al.*, 2016]:

$$\mathbf{Z}_{i,j}^{(v)} = \begin{cases} 0, & i = j \\ \left(\widetilde{\mathbf{Z}}_{i,j}^{(v)} + \eta\right)_+, & i \neq j \end{cases} \tag{12}$$

with $\eta = \frac{1}{n-1} - \frac{1}{n-1}\sum_{j=1,j\neq i}^{N} \widetilde{\mathbf{Z}}_{i,j}^{(v)}$.

(3) $\mathbf{B}^{(v)}$-Step: The solution for $\mathbf{B}^{(v)}$ can be divided into the following two parts for solving:

$$\begin{cases} \min\limits_{\mathbf{B}_1^{(v)}} \sum\limits_{v=1}^{m}\left(\lambda_1\left\|\mathbf{B}_1^{(v)}\right\|_1 + \left\|\mathbf{E}_1^{(v)} - \mathbf{B}_1^{(v)}\right\|_F^2\right) \\ \min\limits_{\mathbf{B}_2^{(v)}} \sum\limits_{v=1}^{m}\left(\lambda_1\left\|\mathbf{B}_2^{(v)}\right\|_1 + \frac{\rho}{2}\left\|\mathbf{E}_2^{(v)} - \mathbf{B}_2^{(v)}\right\|_F^2\right) \end{cases} \tag{13}$$

where $\mathbf{E}_1^{(v)} = \mathbf{X}^{(v)} - \mathbf{X}^{(v)}\mathbf{Z}^{(v)}$ and $\mathbf{E}_2^{(v)} = \mathbf{L}^{(v)} - \mathbf{L}_c + \frac{\mathbf{C}_2^{(v)}}{\rho}$, Eq. (13) is a sparsity constraint optimization problem and has the following two closed-form solutions [Wen *et al.*, 2019]:

$$\begin{cases} \mathbf{B}_1^{(v)} = \vartheta_{\lambda_1}\left(\mathbf{X}^{(v)} - \mathbf{X}^{(v)}\mathbf{Z}^{(v)}\right) \\ \mathbf{B}_2^{(v)} = \vartheta_{\lambda_1/\rho}\left(\mathbf{L}^{(v)} - \mathbf{L}_c + \mathbf{C}_2^{(v)}/\rho\right) \end{cases} \tag{14}$$

where ϑ represents the shrinkage operator.

(4) $\mathbf{L}^{(v)}$-Step: Similar to $\mathbf{X}_M^{(v)}$, we can obtain the updating rule of $\mathbf{L}_v$ by taking the derivative of corresponding minimization problem and setting it to zero as:

$$\mathbf{L}^{(v)} = (\mathbf{D}_1^{(v)} + \mathbf{D}_2^{(v)} + \mathbf{D}_3^{(v)})/3 \tag{15}$$

where $\mathbf{D}_1^{(v)} = \mathbf{Z}^{(v)} - \mathbf{S}^{(v)} + \frac{\mathbf{C}_1^{(v)}}{\rho}$, $\mathbf{D}_2^{(v)} = \mathbf{L}_c + \mathbf{B}_2^{(v)} + \frac{\mathbf{C}_2^{(v)}}{\rho}$, and $\mathbf{D}_3^{(v)} = \mathbf{A}^{(v)} - \frac{\mathbf{Y}^{(v)}}{\rho}$.

(5) $\mathcal{S}$-Step: Removing the unrelated variables, $\mathcal{S}$ can be obtained by optimizing the following problem:

$$\mathcal{S}^* = \arg\min_{\mathcal{S}} \frac{\lambda_2}{\rho}\|\mathcal{S}\|_{2,1} + \frac{1}{2}\|\mathcal{S} - \mathcal{K}\|_F^2 \tag{16}$$

Algorithm 1 CAL Algorithm

Input: Incomplete multi-view data $\{\mathbf{X}_E^{(v)}\}_{v=1}^m$, padding matrix $\{\mathbf{F}_E^{(v)}\}_{v=1}^m$ and $\{\mathbf{F}_M^{(v)}\}_{v=1}^m$, clustering number k;
Parameter: λ_1, λ_2, ρ, η and ρ_{max};
Initialize: Let $\mathbf{Z}^{(v)}$, $\mathbf{B}^{(v)}$, $\mathbf{L}^{(v)}$, $\mathbf{L}_c$, $\mathbf{A}^{(v)}$, $\mathbf{A}_c$, $\mathbf{C}_2^{(v)}$, $\mathcal{S}$, $\mathcal{Y}$ and $\mathcal{C}_1$ to be all zero;
1: **while** not converged **do**
2: Update the variables $\mathbf{X}_M^{(v)}$, $\mathbf{Z}^{(v)}$, $\mathbf{B}^{(v)}$ and $\mathbf{L}^{(v)}$ of each view by Eqs. (8, 12, 14, 15).
3: Update the variables $\mathcal{S}$, $\mathcal{A}$, $\mathbf{L}_c$ and Lagrange multipliers by Eqs. (16, 17, 18, 19).
4: **end while**
Output: $\mathbf{L}^{(v)}$, $\mathbf{L}_c$;
5: Derive R by Eq. 6 and perform spectral clustering on it.

where $\mathcal{K} = \mathcal{Z} - \mathcal{L} + \frac{\mathcal{C}_1}{\rho}$. It can be resolved with the aid of *lemma 4.1* in [Liu *et al.*, 2013].

(6) $\mathcal{A}$-Step: Fixing other variables, the objective function of auxiliary variable $\mathcal{A}$ is formulated as:

$$\mathcal{A}^* = \arg\min_{\mathcal{A}} \|\mathcal{A}\|_{\omega,\mathfrak{Sp}}^p + \frac{\rho}{2}\left\|\mathcal{T} - \mathcal{A} + \frac{\mathcal{Y}}{\rho}\right\|_F^2 \tag{17}$$

It can be resolved using *Theorem 4* in [Gao *et al.*, 2020].

(7) $\mathbf{L}_c^{(v)}$-Step: Similar to the process of solving for $\mathbf{L}^{(v)}$, the optional solution of $\mathbf{L}_c$ is given by

$$L_c = (\rho\mathbf{H}_2^{(v)} + \sum_{v=1}^{m}\rho\mathbf{H}_1^{(v)})/(\rho m + \rho) \tag{18}$$

where $\mathbf{L}^{(v)} - \mathbf{B}_2^{(v)} + \frac{\mathbf{C}_2^{(v)}}{\rho}$ and $\mathbf{H}_2^{(v)} = \mathbf{A}_c^{(v)} - \frac{\mathbf{Y}^{(v+1)}}{\rho}$.

(8) Lagrange multipliers-Step: The updating rules of all Lagrangian multipliers are listed below.

$$\begin{aligned} \mathcal{C}_1 &= \mathcal{C}_1 + \rho\left(\mathcal{Z} - \mathcal{L} - \mathcal{S}\right), \\ C_2^{(v)} &= C_2^{(v)} + \rho\left(L^{(v)} - L_c - B_2^{(v)}\right), \\ \mathcal{Y} &= \mathcal{Y} + \rho\left(\mathcal{T} - \mathcal{A}\right), \end{aligned} \tag{19}$$

where $\rho = \min(\rho\eta,\ \rho_{\max})$. In summary, the whole optimization procedure is presented in Algorithm 1.

3.2 Computational Complexity Analysis

In detail, the complexity of matrix inverse operation for variables $\{\mathbf{X}_M^{(v)}\}_{v=1}^m$ and $\{\mathbf{Z}^{(v)}\}_{v=1}^m$ in each iteration is denoted as $\mathbf{O}(\sum_{v=1}^{m}(N - N_E^{(v)})^3)$ and $\mathbf{O}(mN^3)$ respectively. For updating rotated tensor $\mathcal{A} \in \mathbb{R}^{N\times(m+1)\times N}$, the major computational costs lie in calculating the 3D FFT and the 3D inverse FFT, each of which has a complexity of $O(2N^2(m+1)\log(N))$. Besides, the N SVDs operation of $N \times (m+1)$ matrices in the Fourier domain takes $O(N^2(m+1)^2)$. Considering the other steps involve only fundamental matrix operations, we disregard their computational costs. Thus, the total computational complexity of Algorithm 1 is about $O(t(\sum_{v=1}^{m}(N - N_E^{(v)})^3 + mN^3 + 2N^2(m+1)\log(N) + N^2(m+1)^2))$.

Dataset	Method\Rate	ACC(%)				NMI(%)				Purity(%)			
		0.1	0.3	0.5	0.7	0.1	0.3	0.5	0.7	0.1	0.3	0.5	0.7
BBCSport	EE-R-IMVC [Liu *et al.*, 2021]	57.76	56.31	54.29	47.41	35.45	33.47	31.26	25.95	47.69	45.31	45.16	54.32
	GIMC_FLSD [Wen *et al.*, 2021a]	81.38	76.38	70.00	46.64	73.01	64.43	56.54	22.89	90.00	85.52	80.52	54.83
	IMVTSC-MVI [Wen *et al.*, 2021b]	76.38	74.14	71.90	67.59	74.96	71.81	69.55	51.80	86.72	84.17	83.10	77.41
	TCIMC [Xia *et al.*, 2022]	77.76	75.69	72.41	48.19	74.04	69.44	60.94	29.53	87.76	86.03	79.83	56.64
	SAGF_IMC [Liang *et al.*, 2023]	74.14	73.45	67.84	43.02	66.04	63.11	53.26	24.80	81.72	79.57	71.90	46.72
	LSIMVC [Liu *et al.*, 2023]	80.69	78.79	72.07	38.10	73.57	67.71	58.65	14.53	89.83	87.76	80.17	41.64
	PIMVC [Deng *et al.*, 2023]	79.40	75.34	70.09	55.43	70.62	65.81	56.31	31.96	88.45	84.83	77.67	61.81
	HCLS_CGL [Wen *et al.*, 2023]	92.24	87.07	79.31	50.86	81.81	71.54	61.23	25.06	92.24	88.79	80.17	50.86
	PSIMVC-PC [Li *et al.*, 2024]	81.55	74.91	73.62	72.67	77.71	70.24	69.47	68.63	88.71	85.09	84.91	84.48
	SCSL [Liu *et al.*, 2024b]	65.43	56.90	48.71	44.48	51.49	40.96	29.62	19.70	73.28	64.31	54.57	49.20
	CAL(Ours)	**94.83**	**92.67**	**91.38**	**91.12**	**93.37**	**90.21**	**88.27**	**86.27**	**95.69**	**94.40**	**91.38**	**92.93**
MSRC	EE-R-IMVC [Liu *et al.*, 2021]	83.24	79.91	77.57	73.61	70.63	67.88	65.80	60.39	83.24	80.38	78.14	74.67
	GIMC_FLSD [Wen *et al.*, 2021a]	79.90	79.29	66.57	44.19	71.61	69.01	52.44	32.11	81.24	80.29	67.62	45.71
	IMVTSC-MVI [Wen *et al.*, 2021b]	94.79	93.12	91.43	72.73	89.57	87.21	85.66	67.95	94.79	93.12	91.43	73.48
	TCIMC [Xia *et al.*, 2022]	86.67	90.19	89.01	57.26	77.81	82.35	80.32	49.89	86.67	90.19	89.01	60.72
	SAGF_IMC [Liang *et al.*, 2023]	82.62	75.48	69.14	50.76	72.62	66.10	58.15	39.43	82.62	76.33	70.81	52.19
	LSIMVC [Liu *et al.*, 2023]	73.86	68.48	60.86	42.57	63.29	55.85	45.62	29.64	74.24	68.90	61.38	44.43
	PIMVC [Deng *et al.*, 2023]	87.71	83.24	77.38	63.67	78.33	72.70	65.68	48.43	87.71	73.43	78.48	64.57
	HCLS_CGL [Wen *et al.*, 2023]	83.33	82.38	75.24	40.00	73.00	70.14	61.26	27.57	83.33	82.38	75.24	40.00
	PSIMVC-PC [Li *et al.*, 2024]	83.33	82.43	80.05	65.33	76.54	73.04	71.19	55.51	84.00	83.86	80.81	67.19
	SCSL [Liu *et al.*, 2024b]	72.76	70.24	64.62	44.52	63.91	62.38	53.97	34.67	74.33	72.33	64.48	46.86
	CAL(Ours)	**99.38**	**99.05**	**98.57**	**96.67**	**98.69**	**97.83**	**96.77**	**92.73**	**99.38**	**99.05**	**98.57**	**96.67**
ORL	EE-R-IMVC [Liu *et al.*, 2021]	71.50	55.50	39.00	25.25	81.57	69.09	54.99	44.95	74.00	57.75	40.50	25.75
	GIMC_FLSD [Wen *et al.*, 2021a]	73.20	60.95	51.50	31.20	85.12	73.36	65.03	47.50	76.95	64.05	54.80	34.05
	IMVTSC-MVI [Wen *et al.*, 2021b]	93.75	87.00	84.75	83.50	96.71	92.35	92.28	88.54	94.75	87.25	86.75	85.00
	TCIMC [Xia *et al.*, 2022]	76.85	75.90	75.90	46.20	87.39	86.98	85.13	66.51	79.45	78.95	77.80	49.75
	SAGF_IMC [Liang *et al.*, 2023]	51.25	49.25	49.50	31.50	75.05	74.18	69.61	51.78	58.00	57.00	53.75	34.00
	LSIMVC [Liu *et al.*, 2023]	77.25	69.50	59.00	39.25	86.66	78.53	72.01	57.97	78.75	71.25	61.75	42.25
	PIMVC [Deng *et al.*, 2023]	74.00	73.25	67.75	45.00	84.80	83.25	78.26	62.91	77.75	77.25	70.75	48.25
	HCLS_CGL [Wen *et al.*, 2023]	75.25	74.25	67.25	41.00	84.27	83.56	76.72	63.25	77.50	76.75	70.75	46.00
	PSIMVC-PC [Li *et al.*, 2024]	64.36	63.76	56.24	38.18	79.87	78.63	73.03	61.17	67.08	65.83	59.09	40.88
	SCSL [Liu *et al.*, 2024b]	79.14	73.80	61.33	38.65	89.72	86.34	78.77	59.79	81.91	78.09	67.21	43.69
	CAL(Ours)	**98.53**	**96.63**	**96.23**	**93.75**	**97.52**	**98.54**	**97.56**	**95.68**	**98.53**	**97.23**	**95.69**	**94.05**
BBC	EE-R-IMVC [Liu *et al.*, 2021]	73.68	70.01	66.92	62.67	53.31	50.85	47.76	47.44	74.85	73.36	71.60	71.24
	GIMC_FLSD [Wen *et al.*, 2021a]	80.12	82.54	80.76	60.76	61.87	64.03	57.97	38.38	80.12	82.54	80.76	66.09
	IMVTSC-MVI [Wen *et al.*, 2021b]	77.71	84.03	83.80	76.31	70.45	75.29	75.74	59.67	80.85	84.03	83.80	79.15
	TCIMC [Xia *et al.*, 2022]	86.64	86.13	81.96	72.34	71.43	68.30	59.89	47.60	87.45	86.58	82.39	74.38
	SAGF_IMC [Liang *et al.*, 2023]	82.92	79.02	77.84	60.66	66.88	61.18	56.69	38.20	82.92	79.02	77.84	61.02
	LSIMVC [Liu *et al.*, 2023]	70.07	66.55	59.21	48.63	58.48	55.86	47.97	43.33	75.97	69.08	63.90	54.99
	PIMVC [Deng *et al.*, 2023]	87.81	84.60	83.66	73.87	71.41	66.14	62.73	48.26	87.81	84.60	83.66	74.04
	HCLS_CGL [Wen *et al.*, 2023]	83.07	82.92	78.54	56.50	61.85	58.63	51.92	27.44	83.07	82.92	78.54	56.50
	PSIMVC-PC [Li *et al.*, 2024]	70.74	68.70	68.04	66.67	60.30	61.13	63.15	65.09	76.80	77.34	79.37	80.86
	SCSL [Liu *et al.*, 2024b]	85.12	84.23	82.96	72.69	70.63	66.40	65.62	47.36	92.88	94.09	90.98	79.05
	CAL(Ours)	**98.35**	**97.88**	**97.50**	**95.62**	**94.32**	**94.17**	**91.83**	**87.66**	**98.35**	**97.88**	**97.50**	**95.62**

Table 1: Clustering results (%) w.r.t. three metrics on BBCSport, MSRC, ORL and BBC datasets with different missing rates.

4 Experiment

Dataset	Sample	View	Cluster	Feature
BBCSport	116	4	5	1991, 2063, 2113, 2158
MSRC	210	5	7	24, 576, 512, 256, 254
ORL	400	4	40	4096, 3304, 6750, 1024
BBC	685	4	5	4659, 4633, 4665, 4684
Flower	1360	7	17	1360, 1360, ..., 1360
Caltech101-7	1474	6	7	48, 40, 254, 1984, 512, 928
Handwritten	2000	5	10	240, 76, 216, 47, 64

Table 2: General statistics of datasets.

4.1 Experimental Settings

Datasets: We evaluate our method on seven popular multi-view datasets, including BBCSport, MSRC, ORL, BBC, Caltech101-7, Flower and Handwritten datasets. Table 2 lists the general statistics.

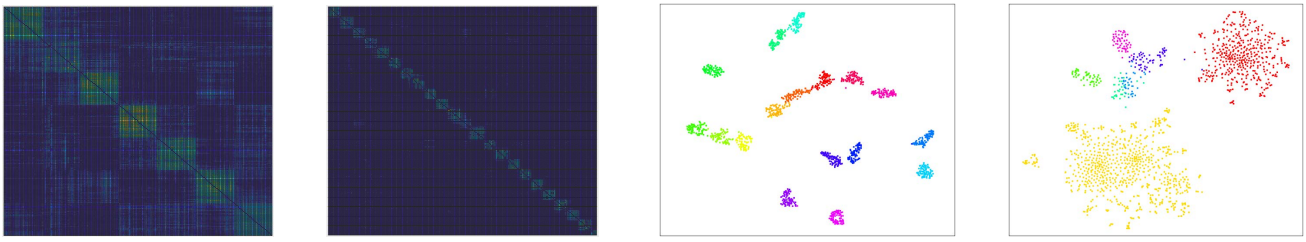
Compared methods: We compare CAL with the following ten state-of-the-art methods: EE-E-IMVC [Liu *et al.*, 2021], GIMC_FLSD [Wen *et al.*, 2021a], IMVTSC-MVI [Wen *et al.*, 2021b], PSIMVC-PC [Li *et al.*, 2024], SAGF_IMC [Liang *et al.*, 2023], TCIMC [Xia *et al.*, 2022], LSIMVC [Liu *et al.*, 2023], PIMVC [Deng *et al.*, 2023], HCLS_CGL [Wen *et al.*, 2023], SCSL [Liu *et al.*, 2024b], among which IMVTSC-MVI and TCIMC are also tensor-based methods.

Incomplete data construction: For each dataset, we arbitrarily remove 10%, 30%, 50%, and 70% samples from each view to construct incomplete multi-view data, while ensuring that each sample exists in at least one view.

Implementation Details: For the comparative methods, we have sought their optimal hyper-parameters following the recommendations in the original papers. For our CAL, we select

Dataset	Method\Rate	ACC(%)				NMI(%)				Purity(%)			
		0.1	0.3	0.5	0.7	0.1	0.3	0.5	0.7	0.1	0.3	0.5	0.7
Flower	EE-R-IMVC [Liu *et al.*, 2021]	50.88	39.56	25.66	16.62	46.00	30.01	16.92	11.15	52.06	40.96	26.99	17.06
	GIMC_FLSD [Wen *et al.*, 2021a]	45.24	41.51	34.74	23.49	46.17	42.49	35.29	24.91	46.63	42.99	35.71	24.63
	IMVTSC-MVI [Wen *et al.*, 2021b]	92.13	87.21	86.47	82.13	87.33	87.45	80.83	80.64	92.91	89.04	86.47	82.13
	TCIMC [Xia *et al.*, 2022]	53.65	50.57	47.37	36.96	50.96	49.35	43.39	33.30	56.47	52.50	49.29	38.19
	SAGF_IMC [Liang *et al.*, 2023]	51.18	49.85	37.87	30.37	51.84	50.27	37.87	29.16	52.79	81.54	39.34	31.40
	LSIMVC [Liu *et al.*, 2023]	50.81	46.32	38.38	32.87	52.37	45.80	38.22	32.55	51.54	47.28	39.04	34.04
	PIMVC [Deng *et al.*, 2023]	62.50	60.81	54.63	49.78	59.72	56.56	50.09	41.66	63.53	62.21	56.25	50.37
	HCLS_CGL [Wen *et al.*, 2023]	47.50	42.79	35.51	28.97	45.19	40.18	32.55	27.31	47.87	42.94	36.40	30.29
	PSIMVC-PC [Li *et al.*, 2024]	37.92	36.27	31.94	25.50	42.22	37.40	32.60	25.76	39.36	37.05	33.32	26.62
	SCSL [Liu *et al.*, 2024b]	46.54	43.96	31.50	27.54	47.07	41.83	32.33	27.60	47.67	44.92	44.92	29.21
	CAL(Ours)	**98.53**	**98.01**	**97.03**	**96.03**	**97.52**	**96.67**	**95.78**	**93.81**	**98.53**	**98.01**	**97.03**	**96.03**
Caltech101-7	EE-R-IMVC [Liu *et al.*, 2021]	41.66	40.37	35.07	27.95	28.33	19.62	16.61	8.79	75.78	71.10	66.01	63.09
	GIMC_FLSD [Wen *et al.*, 2021a]	58.56	52.24	52.09	46.59	14.95	16.06	17.95	18.81	66.24	65.40	67.34	68.70
	IMVTSC-MVI [Wen *et al.*, 2021b]	63.37	62.48	61.94	46.34	54.39	53.02	51.67	32.42	87.38	86.91	87.52	79.17
	TCIMC [Xia *et al.*, 2022]	64.86	64.03	61.68	60.28	52.40	46.91	47.28	41.38	86.23	84.82	83.60	82.56
	SAGF_IMC [Liang *et al.*, 2023]	71.64	66.83	62.55	56.78	64.59	58.67	49.25	36.25	87.99	86.09	83.04	79.38
	LSIMVC [Liu *et al.*, 2023]	71.44	70.28	63.70	60.31	61.12	51.82	50.01	35.72	90.77	84.33	84.74	80.05
	PIMVC [Deng *et al.*, 2023]	70.62	66.42	66.28	65.33	56.34	52.43	52.17	48.94	88.60	85.82	86.16	85.07
	HCLS_CGL [Wen *et al.*, 2023]	**72.66**	**67.77**	65.26	63.98	51.95	50.97	52.16	45.63	84.94	86.43	84.46	83.99
	PSIMVC-PC [Li *et al.*, 2024]	50.21	50.00	48.37	49.29	47.26	42.96	39.48	30.52	82.63	80.60	79.17	74.91
	SCSL [Liu *et al.*, 2024b]	66.35	58.21	56.93	56.38	50.22	50.95	21.85	16.11	82.23	83.12	70.43	66.76
	CAL(Ours)	71.78	67.39	**67.03**	**65.67**	**70.06**	**66.13**	**66.14**	**64.27**	**95.12**	**94.89**	**94.91**	**94.37**
Handwritten	EE-R-IMVC [Liu *et al.*, 2021]	88.60	85.23	76.70	51.74	78.64	73.30	62.26	40.21	88.60	85.23	76.70	51.74
	GIMC_FLSD [Wen *et al.*, 2021a]	88.65	87.89	86.85	86.35	80.01	78.44	76.23	75.33	88.65	87.45	86.56	86.55
	IMVTSC-MVI [Wen *et al.*, 2021b]	96.43	95.21	91.83	87.15	92.44	89.95	84.07	77.39	96.43	95.21	91.83	87.15
	TCIMC [Xia *et al.*, 2022]	83.33	83.72	82.69	84.15	86.39	85.08	79.89	80.88	85.93	85.85	82.97	84.27
	SAGF_IMC [Liang *et al.*, 2023]	84.40	84.12	83.61	81.19	89.49	87.71	85.63	73.06	86.99	86.42	85.70	82.01
	LSIMVC [Liu *et al.*, 2023]	91.97	93.63	87.42	85.75	85.63	88.01	82.75	76.46	91.97	96.63	88.10	85.75
	PIMVC [Deng *et al.*, 2023]	94.80	93.79	91.23	88.61	89.74	87.71	83.88	79.56	94.88	93.79	91.23	88.61
	HCLS_CGL [Wen *et al.*, 2023]	85.60	85.05	81.35	81.00	88.50	88.18	81.26	76.96	87.70	87.50	83.40	81.50
	PSIMVC-PC [Li *et al.*, 2024]	86.79	85.15	79.12	75.09	83.08	81.61	77.01	69.50	87.09	85.63	79.83	75.47
	SCSL [Liu *et al.*, 2024b]	94.66	92.44	85.30	80.53	88.28	82.67	78.15	78.82	96.89	94.06	87.07	82.96
	CAL(Ours)	**100.00**	**100.00**	**99.95**	**99.25**	**99.95**	**99.95**	**99.86**	**98.24**	**99.95**	**99.95**	**99.95**	**99.25**

Table 3: Clustering results (%) w.r.t. three metrics on Flower, Caltech101-7 and Handwritten datasets with different missing rates.



(a) MSRC (b) ORL (c) Flower (d) Caltech01-7

Figure 3: Visualization of the unified structural graph on MSRC, ORL, Flower, and Caltech101-7 datasets with a 50% missing rate.

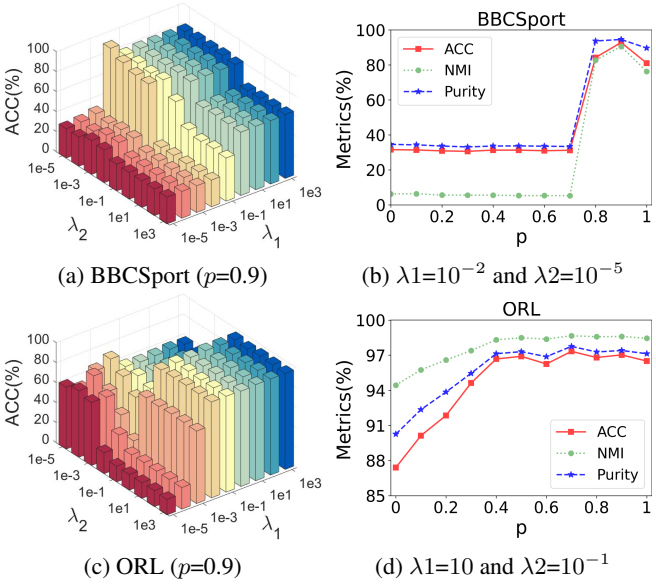
$\lambda 1$ and $\lambda 2$ within the range of $[10^{-5}, 10^{3}]$ and evaluate our method using three popular metrics: accuracy (ACC), normalized mutual information (NMI), and purity.

4.2 Experimental Results and Analysis

All experimental results on seven datasets w.r.t. three metrics are illustrated in table 1 and table 3, where the best results are in **red** and the second-best are highlighted in blue. From the experimental results, we can reach the following observations: (1) Our method almost achieves the best performance on all datasets and exhibits significant improvements in cases of high missing rates. For instance, on BBCSport dataset with a missing rate of 70%, our algorithm achieves advancement of 18.45%-53.02%, 17.64%-71.74%, and 8.45%-51.29% in terms of three respective metrics. (2) Compared with the second-best performing method (IMVTSC-MVI) on most datasets, our method shows a substantial enhancement in the NMI metric under equal or lower ACC, which not only illustrates the superiority of high-order interaction across multiple views over dual-view interaction but also validates that consensus guidance can boost the ability to learn consistency. (3) We employ heatmaps and t-SNE to visualize the unified structural graph R and show in Figure 3, that the structural graph exhibits a distinct block diagonal structure and compact clusters, indicating that our method more effectively explores the data clustering structure. (4) From the perspective of interaction mechanisms, CAL establishes

Dataset	Method\Rate	ACC(%)				NMI(%)				Purity(%)			
		0.1	0.3	0.5	0.7	0.1	0.3	0.5	0.7	0.1	0.3	0.5	0.7
BBCSport	CAL-I	88.45	86.38	83.62	76.81	74.28	72.13	70.21	66.16	88.45	87.24	84.83	81.98
	CAL-II	90.17	87.24	85.17	81.64	86.82	80.16	76.27	62.60	93.28	90.69	90.43	82.33
	CAL	**94.83**	**92.67**	**91.38**	**91.12**	**93.37**	**90.21**	**88.27**	**86.27**	**95.69**	**94.40**	**91.38**	**92.93**
ORL	CAL-I	93.28	93.10	92.78	91.63	97.06	96.88	95.53	93.39	94.43	94.15	93.35	91.82
	CAL-II	90.65	90.25	89.95	78.28	95.26	94.47	93.76	86.23	91.85	91.13	90.88	79.95
	CAL	**98.53**	**96.63**	**96.23**	**93.75**	**97.52**	**98.54**	**97.56**	**95.68**	**98.53**	**97.23**	**95.69**	**94.05**

Table 4: The ablation experiments on BBCSport and ORL datasets with different missing rates.



(a) BBCSport (p=0.9) (b) $\lambda1$=10^{-2} and $\lambda2$=10^{-5}

(c) ORL (p=0.9) (d) $\lambda1$=10 and $\lambda2$=10^{-1}

Figure 4: Parameter analysis of CAL w.r.t. $\lambda1$, $\lambda2$ and p on BBCSport and ORL datasets with a 30% missing rate.

consensus-aware interactions based on existing tensor rank constraints compared with other methods, which can optimize the interaction process through consensus structure and achieve dynamic fusion between specific and consensus representation. Additionally, CAL can retain complete and accurate data clustering relationships through the dual affinity constraint mechanism composed of sample-wise affinities learning and intrinsic representation learning.

Parameter sensitivity analysis: In this subsection, we analyze the impact of model parameters $\lambda1$ and $\lambda2$ on clustering performance. Figure 4(a) and (c) illustrates the variation in ACC for different combinations of $\lambda1$ and $\lambda2$ on BBCSport and ORL datasets with a 30% missing rate. It can be observed that satisfactory performance is achieved when $\lambda1$ varies within the range $[10^{-2}, 10^{3}]$. The clustering results for BBCSport and ORL datasets are insensitive to relatively large and small values of $\lambda2$, respectively. Furthermore, we also investigate the influence of power parameter p in Figure 4(b) and (d). Generally, CAL obtains ideal clustering results when p is selected from [0.8, 0.9].

Convergence analysis: This paper develops an alternative optimization algorithm for solving the proposed CAL model. By iteratively optimizing each sub-problem and obtaining a closed-form solution, we ensure that the algorithm is bounded below. Additionally, convergence curves on Flower and Caltech101-7 datasets with a 30% missing rate are plotted in Figure 5, depicting the evolution of algorithm loss and ACC throughout iterations. It is evident that the algorithm loss declines promptly and has a lower bound of 0.

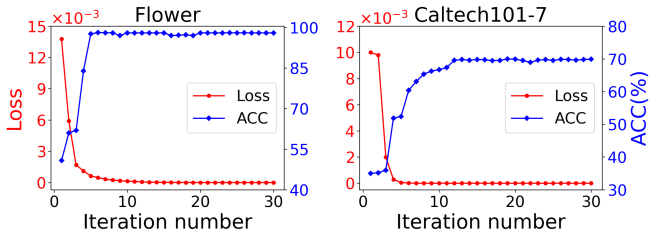


Figure 5: Convergence curves on Flower and Caltech101-7 datasets with a 30% missing rate.

Ablation study: We conduct the ablation study to validate the effectiveness of each submodel and record the results on BBCSport and ORL datasets ith different missing rates. In specific, based on Eq. (5), we establish the below two degenerate methods: CAL without structured sparsity constraint (CAL-I) and high-order interaction without consensus guidance (CAL-II). From the results in table 4, it is evident that the clustering results of both variants are inferior to the original method. Moreover, in most cases, the performance degradation of CAL-II from the missing rate of 0.1 to 0.7 across the three metrics is greater than that of CAL-I. Large fluctuations in the performance of CAL-II underscores the substantial role of consensus representation in advancing consistency learning among multiple views in high-order interaction.

5 Conclusion

In this paper, we propose a cross-view affinity learning method named CAL for incomplete multi-view clustering, which effectively avoids the limitations posed by dual-view interaction on information fusion. CAL integrates comprehensive content information and pure structural information to acquire a unified structural graph for clustering under the guidance of consensus representation through the high-order interaction across multiple views. Moreover, an effective optimization algorithm for the CAL has been developed and extensive experiments have been carried out to validate its superiority over state-of-the-art methods.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant U21A20470, Grant 62172136 and Grant 72188101, by National Key Research and Development Program of China Grant 2024YFB4710800, by the Institute of Advanced Medicine and Frontier Technology under Grant 2023 IHM01080, Liaoning Provincial Natural Science Foundation Grant 2024-MS-012, Dalian Science and Technology Talent Innovation Support Plan, Grant 2024RY010.

References

[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 (CVPR)*, June 2015.

[Chen *et al.*, 2025] Yawei Chen, Huibing Wang, Jinjia Peng, and Yang Wang. Anchor learning with potential cluster constraints for multi-view clustering. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(15):15939–15947, Apr. 2025.

[Deng *et al.*, 2023] Shijie Deng, Jie Wen, Chengliang Liu, Ke Yan, Gehui Xu, and Yong Xu. Projective incomplete multi-view clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[Fang *et al.*, 2023] Uno Fang, Man Li, Jianxin Li, Longxiang Gao, Tao Jia, and Yanchun Zhang. A comprehensive survey on multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 35(12):12350–12368, 2023.

[Gao *et al.*, 2020] Quanxue Gao, Pu Zhang, Wei Xia, Deyan Xie, Xinbo Gao, and Dacheng Tao. Enhanced tensor rpca and its application. *IEEE transactions on pattern analysis and machine intelligence*, 43(6):2133–2140, 2020.

[Gu *et al.*, 2024] Zhibin Gu, Zhendong Li, and Songhe Feng. EDISON: Enhanced dictionary-induced tensorized incomplete multi-view clustering with gaussian error rank minimization. In *Forty-first International Conference on Machine Learning*, 2024.

[Huang *et al.*, 2022] Shudong Huang, Hongjie Wu, Yazhou Ren, Ivor Tsang, Zenglin Xu, Wentao Feng, and Jiancheng Lv. Multi-view subspace clustering on topological manifold. In *Advances in Neural Information Processing Systems*, volume 35, pages 25883–25894, 2022.

[Kang *et al.*, 2020] Zhao Kang, Wangtao Zhou, Zhitong Zhao, Junming Shao, Meng Han, and Zenglin Xu. Large-scale multi-view subspace clustering in linear time. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(04):4412–4419, 2020.

[Li *et al.*, 2022] Zhenglai Li, Chang Tang, Xiao Zheng, Xinwang Liu, Wei Zhang, and En Zhu. High-order correlation preserved incomplete multi-view subspace clustering. *IEEE Transactions on Image Processing*, 31:2067–2080, 2022.

[Li *et al.*, 2024] Miaomiao Li, Siwei Wang, Xinwang Liu, and Suyuan Liu. Parameter-free and scalable incomplete multiview clustering with prototype graph. *IEEE Transactions on Neural Networks and Learning Systems*, 35(1):300–310, 2024.

[Liang *et al.*, 2023] Naiyao Liang, Zuyuan Yang, and Shengli Xie. Incomplete multi-view clustering with sample-level auto-weighted graph fusion. *IEEE Transactions on Knowledge and Data Engineering*, 35(6):6504–6511, 2023.

[Liang *et al.*, 2024] Cheng Liang, Daoyuan Wang, Huaxiang Zhang, Shichao Zhang, and Fei Guo. Robust tensor subspace learning for incomplete multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 36(11):6934–6948, 2024.

[Liu *et al.*, 2013] Guangcan Liu, Zhouchen Lin, Shuicheng Yan, Ju Sun, Yong Yu, and Yi Ma. Robust recovery of subspace structures by low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):171–184, 2013.

[Liu *et al.*, 2021] Xinwang Liu, Miaomiao Li, Chang Tang, Jingyuan Xia, Jian Xiong, Li Liu, Marius Kloft, and En Zhu. Efficient and effective regularized incomplete multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(8):2634–2646, 2021.

[Liu *et al.*, 2023] Chengliang Liu, Zhihao Wu, Jie Wen, Yong Xu, and Chao Huang. Localized sparse incomplete multi-view clustering. *IEEE Transactions on Multimedia*, 25:5539–5551, 2023.

[Liu *et al.*, 2024a] Cheng Liu, Rui Li, Hangjun Che, Man-Fai Leung, Si Wu, Zhiwen Yu, and Hau-San Wong. Latent structure-aware view recovery for incomplete multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 36(12):8655–8669, 2024.

[Liu *et al.*, 2024b] Suyuan Liu, Junpu Zhang, Yi Wen, Xihong Yang, Siwei Wang, Yi Zhang, En Zhu, Chang Tang, Long Zhao, and Xinwang Liu. Sample-level cross-view similarity learning for incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 14017–14025, 2024.

[Long *et al.*, 2024a] Zhen Long, Qiyuan Wang, Yazhou Ren, Yipeng Liu, and Ce Zhu. S2mvtc: a simple yet efficient scalable multi-view tensor clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 26213–26222, June 2024.

[Long *et al.*, 2024b] Zhen Long, Ce Zhu, Pierre Comon, Yazhou Ren, and Yipeng Liu. Feature space recovery for efficient incomplete multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 36(9):4664–4677, 2024.

[Nie *et al.*, 2016] Feiping Nie, Xiaoqian Wang, Michael Jordan, and Heng Huang. The constrained laplacian rank algorithm for graph-based clustering. *Proceedings of the AAAI Conference on Artificial Intelligence*, 30(1), 2016.

[Su *et al.*, 2024] Peng Su, Yixi Liu, Shujian Li, Shudong Huang, and Jiancheng Lv. Robust contrastive multi-view kernel clustering. In Kate Larson, editor, *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pages 4938–4945. International Joint Conferences on Artificial Intelligence Organization, 8 2024. Main Track.

[Tan *et al.*, 2024] Yuze Tan, Hecheng Cai, Shudong Huang, Shuping Wei, Fan Yang, and Jiancheng Lv. An effective augmented lagrangian method for fine-grained multiview optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 15258–15266, 2024.

[Wang *et al.*, 2023] Huibing Wang, Guangqi Jiang, Jinjia Peng, Ruoxi Deng, and Xianping Fu. Towards adaptive consensus graph: Multi-view clustering via graph collaboration. *IEEE Transactions on Multimedia*, 25:6629–6641, 2023.

[Wang *et al.*, 2024a] Huibing Wang, Mingze Yao, Yawei Chen, Yunqiu Xu, Haipeng Liu, Wei Jia, Xianping Fu, and Yang Wang. Manifold-based incomplete multi-view clustering via bi-consistency guidance. *IEEE Transactions on Multimedia*, 26:10001–10014, 2024.

[Wang *et al.*, 2024b] Huibing Wang, Mingze Yao, Guangqi Jiang, Zetian Mi, and Xianping Fu. Graph-collaborated auto-encoder hashing for multiview binary clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 35(7):10121–10133, 2024.

[Wang, 2021] Yang Wang. Survey on deep multi-modal data analytics: Collaboration, rivalry, and fusion. *ACM Trans. Multimedia Comput. Commun. Appl.*, 17(1s), March 2021.

[Wen *et al.*, 2019] Jie Wen, Xiaozhao Fang, Jinrong Cui, Lunke Fei, Ke Yan, Yan Chen, and Yong Xu. Robust sparse linear discriminant analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(2), 2019.

[Wen *et al.*, 2020] Jie Wen, Ke Yan, Zheng Zhang, Yong Xu, Junqian Wang, Lunke Fei, and Bob Zhang. Adaptive graph completion based incomplete multi-view clustering. *IEEE Transactions on Multimedia*, 23:2493–2504, 2020.

[Wen *et al.*, 2021a] Jie Wen, Zheng Zhang, Zhao Zhang, Lunke Fei, and Meng Wang. Generalized incomplete multiview clustering with flexible locality structure diffusion. *IEEE Transactions on Cybernetics*, 51(1):101–114, 2021.

[Wen *et al.*, 2021b] Jie Wen, Zheng Zhang, Zhao Zhang, Lei Zhu, Lunke Fei, Bob Zhang, and Yong Xu. Unified tensor framework for incomplete multi-view clustering and missing-view inferring. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 10273–10281, 2021.

[Wen *et al.*, 2023] Jie Wen, Chengliang Liu, Gehui Xu, Zhihao Wu, Chao Huang, Lunke Fei, and Yong Xu. Highly confident local structure based consensus graph learning for incomplete multi-view clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15712–15721, 2023.

[Wu *et al.*, 2024] Tingting Wu, Songhe Feng, and Jiazheng Yuan. Low-rank kernel tensor learning for incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 15952–15960, 2024.

[Xia *et al.*, 2022] Wei Xia, Quanxue Gao, Qianqian Wang, and Xinbo Gao. Tensor completion-based incomplete multiview clustering. *IEEE Transactions on Cybernetics*, 52(12):13635–13644, 2022.

[Yao *et al.*, 2025] Mingze Yao, Huibing Wang, Yawei Chen, and Xianping Fu. Between/within view information completing for tensorial incomplete multi-view clustering. *IEEE Transactions on Multimedia*, 2025.

[Zhao *et al.*, 2023] Shuping Zhao, Jie Wen, Lunke Fei, and Bob Zhang. Tensorized incomplete multi-view clustering with intrinsic graph completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 11327–11335, 2023.