ALIGNING COLLABORATIVE VIEW RECOVERY AND TENSORIAL SUBSPACE LEARNING VIA LATENT REPRESENTATION FOR INCOMPLETE MULTI-VIEW CLUSTERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-view data usually suffer from partially missing views in open scenarios, which inevitably degrades clustering performance. The incomplete multi-view clustering (IMVC) has attracted increasing attention and achieved significant success. Although existing imputation-based IMVC methods perform well, they still face one crucial limitation, i.e., view recovery and subspace representation lack explicit alignment and collaborative interaction in exploring complementarity and consistency across multiple views. To this end, this study proposes a novel IMVC method to **A**lign collaborative view **R**ecovery and tensorial **S**ubspace **L**earning via latent representation (ARSL-IMVC). Specifically, the ARSL-IMVC infers the complete view from view-shared latent representation and view-specific estimator with Hilbert-Schmidt Independence Criterion regularizer, reshaping the consistent and diverse information intrinsically embedded in original multi-view data. Then, the ARSL-IMVC learns the view-shared and view-specific subspace representations from latent feature and recovered views, and models high-order correlations at the global and local levels in the unified low-rank tensor space. Thus, leveraging the latent representation as a bridge in a unified framework, the ARSL-IMVC seamlessly aligns the complementarity and consistency exploration across view recovery and subspace representation learning, negotiating with each other to promote clustering. Extensive experimental results on seven datasets demonstrate the powerful capacity of ARSL-IMVC in complex incomplete multi-view clustering tasks under various view missing scenarios.

1 INTRODUCTION

Recently, the rapid development of information technology promotes the complexity of data forms, including text, audio, images, etc. Extracting available information from multi-view data collected from different sources is challenging, especially in unsupervised scenarios (Xu et al., 2013). As a key unsupervised learning technology, clustering is also expanded to multi-view clustering to better accommodate complex multi-view data (Chao et al., 2021; Chen et al., 2022). Multi-view clustering (MVC) aims to divide unlabeled multi-view data into disjoint clusters by leveraging the consistency and complementarity across multi-view data. Among various MVC approaches, multi-view subspace clustering (Gao et al., 2015; Cao et al., 2015; Shi et al., 2024) receives considerable attention due to its significant performance advantage and robustness. It aims to learn optimal subspace representations in which the samples can be more clearly separated into distinct clusters, simultaneously preserving the complementary information and mitigating redundancy and noise across multiple views. When both inter-view consistency and view-specific diversity are fully explored and well balanced, the clustering performance can be significantly enhanced (Guo et al., 2023).

However, in open scenarios, due to sensor failure, missing annotations, or data corruption, etc, it is usually difficult to obtain complete data for all views. Reducing the negative impact of incompleteness of multiple views on clustering performance becomes a major challenge currently faced in the MVC filed (Wen et al., 2023b). And thus, the incomplete multi-view clustering (IMVC) methods are widely proposed and divided into two categories, i.e., imputation-based and imputation-free

methods. The imputation-based IMVC methods adhere to the mechanism of first view recovery and then clustering structure exploration (Liu et al., 2024a; Wang et al., 2021). The imputation-free IMVC methods only focus on partial observable views to explore clustering information, avoiding computational costs of missing view completion (Wen et al., 2024; Qin et al., 2025). Despite strong simplicity of imputation-free IMVC methods, their discriminability to extract clustering information is restricted by limited available views, especially when the missing rate is high.

The imputation-based IMVC methods utilize heuristic or learnable strategies to impute the missing views, providing strong data foundation for clustering information exploration and enhancing the interpretability. Surprisingly, several methods jointly recover the missing views and learn clustering representations (subspace coefficient, graph similarity, cluster indicator, etc) in a unified framework, significantly improving clustering quality. Despite the success, two critical challenges remain in most existing imputation-based IMVC methods. First, the recovered or completed views often suffer from limited structural fidelity and insufficient diversity and consistency reshaping, which are both essential for effective multi-view clustering. More important, there is no explicit alignment and collaborative interaction in view recovery and subspace representation learning in exploring complementarity and consistency.

To this end, this study unifies collaborative view completion and tensorial subspace learning, and breaks the gap between them in complementarity and consistency modeling by shared latent representation. Notably, the latent representation not only serves as fictitious transitional factor for view reconstruction but also directly contributes to the subspace learning with structural awareness. Along with view-specific diversity term, the proposed method provides more freedom in enriching feature description across views. Consequently, the shared and specific subspace representations derived from the latent space and imputed views are integrated into low-rank tensor, enabling interactions across different levels of structural information. In summary, the primary contributions of this study are as follows:

- The proposed novel IMVC method facilitates the unified and explicit alignment of complementarity and consistency exploration across both missing feature reconstruction and subspace representation learning, fostering a coherent cross-view correlation modeling.
- The complex high-order correlation among local specific and global shared subspace representations is collaboratively explored. And the structural semantics embedded in subspace representations are fed back to latent representation and recovered views, improving the imputation fidelity and clustering discriminability.
- An effective iterative method is designed to solve the optimization problem. Extensive experiments verify the superiority of proposed method.

2 Related Work

Recently, numerous IMVC methods have been widely proposed (Wen et al., 2020a; Shen et al., 2025; Jiang et al., 2025; Li et al., 2024), which could be roughly grouped into two categories according to the way of handling missing samples. The first category crudely ignores missing samples and focus on learning clustering representation from available views, i.e., imputation-free. Li et al. divided samples into view-complete parts and view-specific missing parts and learned low-dimensional representation by non-negative matrix factorization (Li et al., 2014). Hu et al. proposed doubly aligned incomplete multi-view clustering (DAIMC), aligning the available views to learn a compact representation shared by all views and trying to weaken the influence of missing samples (Hu & Chen, 2018). Due to the powerful relationship representation ability of graph, the incompleteness of data is transferred to the similarity domain (Wen et al., 2020b; 2023a). Wen et al. designed a graph-based IMVC method, Incomplete Multi-view Spectral Clustering with Adaptive Graph Learning (IMSC-AGC), where partial graph for each view is adaptively constructed by only leveraging observable samples and expanded to complete graph, and then a shared spectral embedding is learned (Wen et al., 2020a). To weaken the impact of noisy view, Wen et al further designed a highly confident local structure induced consensus graph learning (HCLS-CGL) for IMVC (Wen et al., 2023a). Some methods also tried to maximize the utilization of available information by simultaneously considering feature and graph structure and establishing a connection between them (Bai et al., 2024; Liu et al., 2024b).

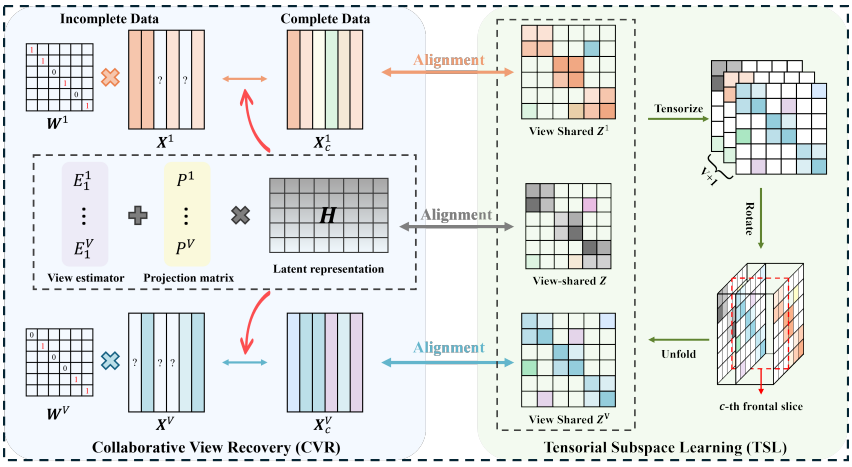


Figure 1: The overall framework of proposed ARSL-IMVC method, which mainly consists of CVR and TSL modules and aligns them in cross-view consistency and complementarity exploration by a latent representation.

Although achieving good progress, the imputation-free methods ignore the underlying connections between samples or views, which instinctively undermine the structural integrity and compromising the clustering performance. To address this issue, imputation-based IMVC methods recover missing data and learn clustering representation on the filling data. Wen et al. proposed a unified embedding alignment framework (UEAF) to infer missing samples, fully leveraging the cross-view correlations by common representation (Wen et al., 2019). Wang et al. leveraged both local view-specific and global view-shared similarity structures to guide the recovery (Wang et al., 2024). To utilize high-order relationships between samples, Chen et al. developed hypergraph induced missing reconstruction strategy for IMVC (IMVC-HG) and further explored the complementarity embedded in inter-view label representations (Chen et al., 2025). Guo et al. proposed robust mixed-order graph learning (RMoGL), utilizing the mixed-order structural information of recovered data from low-order to high-order (Guo et al., 2025). In addition, several methods focus on modeling both complex cross-view correlation and similarity relationships between samples by low-rank tensor learning (Li et al., 2022; Yao et al., 2025; Chen et al., 2024). Li et al. explored the high-order correlation of views and that of samples, utilizing hyper-Laplacian regularizer in view recovery and low-rank tensor learning in subspace structure recovery (Li et al., 2022). Yao et al. focused on recovering the multi-view similarity graphs and utilized dual tensor constraint to explore complex correlations, which proposes the between/within view information completing for tensorial incomplete multi-view clustering (BWIC-TIMC) (Yao et al., 2025). In addition, to improve data separability, some kernel-based IMVC methods reconstruct the missing data in kernel Hilbert space (Liu et al., 2020; Wu et al., 2024; Liu et al., 2021; Wu et al., 2025). Although most existing IMVC methods unifying missing view recovery and clustering representation learning have achieved notable progress, they are only treated as weakly coupled strategies. Specifically, they fail to simultaneously exploit the inherent consistency and complementarity among multiple views during the view recovery and representation learning and achieve the semantic correlation exploration alignment through explicit bridge, resulting in shallow collaborative interaction between them and sub-optimal clustering structure. Thus, this study explicitly aligns the collaborative view recovery and tensorial subspace learning by a shared latent representation for IMVC (ARSL-IMVC), bridging the tight interaction between them.

3 Methodology

In this section, the proposed ARSL-IMVC model is illustrated in detail from two main modules: Collaborative View Recovery (CVR) and Tensorial Subspace Learning (TSL), where the framework is presented in Figure 1. Then, the iterative optimization procedure is provided. The basic notations are conveniently summarized in the Table 1.

Table 1: Main notations and descriptions in this study.

Notations	Descriptions	Notations	Descriptions
$\boldsymbol{X}^v \in \mathbb{R}^{d_v \times n}$	Data matrix of the v-th view	$\boldsymbol{E}_1^v$	View recovery estimator
$\boldsymbol{W}^v \in \{0,1\}^{n \times n}$	Missing indicator matrix	$\boldsymbol{E}_2^v, \boldsymbol{E}_H$	Noise matrices
$\boldsymbol{P}^v \in \mathbb{R}^{d_v \times k}$	Projection matrix of the v-th view	$\boldsymbol{Z}^v, \boldsymbol{Z} \in \mathbb{R}^{n \times n}$	Specific and shared representations
$\boldsymbol{H} \in \mathbb{R}^{k \times n}$	Latent representation matrix	$\mathcal{Z} \in \mathbb{R}^{n \times (V+1) \times n}$	Subspace representation tensor

3.1 Collaborative View Recovery

Consider an incomplete multi-view dataset consisting of n samples from V views, i.e., $\{\{\boldsymbol{x}_i^v\}_{i=1}^n\}_{v=1}^V$, where $\boldsymbol{x}_i^v \in \mathbb{R}^{d_v}$ is the i-th sample in v-th view. The incomplete pattern is symbolized with a diagonal indicator matrix $\boldsymbol{W}^v \in \{0,1\}^{n \times n}$ such that $\boldsymbol{W}_{ii}^v = 1$ if i-th sample exists in v-th view and 0 otherwise. The MVC adheres to a convincing assumption that multi-view data is usually embedded in the shared latent space. With this inverse assumption, the ARSL-IMVC tries to linearly infer missing views from a "virtua" latent representation $\boldsymbol{H} \in \mathbb{R}^{k \times n}$ via reconstruction operator $\boldsymbol{P}^v \in \mathbb{R}^{d_v \times k}$, which is beneficial for reshaping the consistency among multiple views. To improve the freedom of view recovery, view-specific feature estimator $\boldsymbol{E}_1^v$ is introduced and thus the view reconstruction is formulated, i.e.,

$$\boldsymbol{P}^v \boldsymbol{H} + \boldsymbol{E}_1^v \tag{1}$$

Further, to ensure that the various recovered views retain sufficient complementary information, the Hilbert-Schmidt Independence Criterion (HSIC) is introduced as diversity regularizer between any estimator pair $\boldsymbol{E}_1^v$ and $\boldsymbol{E}_1^w$, where the empirical HSIC term is defined (Gretton et al., 2007):

$$\text{HSIC}(\boldsymbol{E}_1^v, \boldsymbol{E}_1^w) = \text{Tr}(\boldsymbol{K}_v \tilde{\boldsymbol{H}} \boldsymbol{K}_w \tilde{\boldsymbol{H}})/(n-1)^2 \tag{2}$$

where $\boldsymbol{K}_v$ and $\boldsymbol{K}_w$ are inner product kernel matrices of $\boldsymbol{E}_1^v$ and $\boldsymbol{E}_1^w$ respectively, $\tilde{\boldsymbol{H}} = \boldsymbol{I}_n - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T$ is the centralized matrix. The HSIC term penalize the dependence between the various reconstruction views, thereby encouraging diversity and informativeness among them in the feature level. To explore the consistency-diversity and ensure the reconstruction quality, the CVR module is formulated as follows:

$$\begin{aligned} \min_{\boldsymbol{H}, \boldsymbol{P}^v, \boldsymbol{E}_1^v} & \sum_{w=1; w \neq v}^{V} \text{HSIC}(\boldsymbol{E}_1^v, \boldsymbol{E}_1^w) \\ \text{s.t. } & \boldsymbol{X}^v \boldsymbol{W}^v = (\boldsymbol{P}^v \boldsymbol{H} + \boldsymbol{E}_1^v) \boldsymbol{W}^v, (\boldsymbol{P}^v)^T \boldsymbol{P}^v = \boldsymbol{I} \end{aligned} \tag{3}$$

where $\boldsymbol{X}^v = [\boldsymbol{x}_1^v, \boldsymbol{x}_2^v, \cdots, \boldsymbol{x}_n^v]$, $\boldsymbol{X}^v \boldsymbol{W}^v$ denotes the non-missing data, and the equality constraint is enforced to ensure reconstruction fidelity of observable samples.

3.2 Tensorial Subspace Learning

Subspace representation learning is an effective method for exploring clustering semantics in low-dimensional embedding space, especially for self-representation based methods. Leveraging the shared latent representation and recovered view-specific features, the global semantics and local clustering structures could be effectively characterized, i.e.,

$$\boldsymbol{H} = \boldsymbol{H}\boldsymbol{Z} + \boldsymbol{E}_H, \boldsymbol{P}^v \boldsymbol{H} + \boldsymbol{E}_1^v = (\boldsymbol{P}^v \boldsymbol{H} + \boldsymbol{E}_1^v)\boldsymbol{Z}^v + \boldsymbol{E}_2^v \tag{4}$$

where $\boldsymbol{Z}$, $\boldsymbol{Z}^v$ are view-shared and view-specific subspace representations, respectively, encoding the comprehensive structural semantics; and $\boldsymbol{E}_H$ and $\boldsymbol{E}_2^v$ are noise terms.

To explore the consistency and complementarity across views in subspace representation level, both shared and specific subspace representations into a unified low-rank tensor. Thus, the TSL module is formulated as follows:

$$\begin{aligned} \min_{\boldsymbol{Z}, \boldsymbol{Z}^v, \boldsymbol{E}_H, \boldsymbol{E}_2^v} & \|\mathcal{Z}\|_{\circledast} + \lambda_1(\|\boldsymbol{E}_H\|_{2,1} + \sum_{v=1}^{V} \|\boldsymbol{E}_2^v\|_{2,1}) \\ \text{s.t. } & \boldsymbol{H} = \boldsymbol{H}\boldsymbol{Z} + \boldsymbol{E}_H, \boldsymbol{P}^v \boldsymbol{H} + \boldsymbol{E}_1^v = (\boldsymbol{P}^v \boldsymbol{H} + \boldsymbol{E}_1^v)\boldsymbol{Z}^v + \boldsymbol{E}_2^v, \\ & \mathcal{Z} = \Phi(\boldsymbol{Z}^1, \boldsymbol{Z}^2, \cdots, \boldsymbol{Z}^V, \boldsymbol{Z}) \end{aligned} \tag{5}$$

where $\Phi(\cdot)$ is a tensor construction function, which stacks the representations $(\{\boldsymbol{Z}^v\}_{v=1}^V, \boldsymbol{Z})$ and rotates to form a subspace representation tensor $\mathcal{Z} \in \mathbb{R}^{n\times(V+1)\times n}$. In the low-rank tensor space, the high-order cross-view correlations at different levels be effectively captured, facilitating semantic alignment across diverse views and achieving collaborative interaction between local and global structural information.

3.3 Formulation of the Proposed ARSL-IMVC

By incorporating the CVR and TSL into a unified framework, the final objective of ARSL-IMVC is formulated as:

$$\begin{aligned}\min_{\Upsilon} \|\mathcal{Z}\|_{\circledast} + \lambda_1(\|\boldsymbol{E}_H\|_{2,1} + \sum_{v=1}^{V}\|\boldsymbol{E}_2^v\|_{2,1}) + \lambda_2 \sum_{w=1;w\neq v}^{V} \text{HSIC}(\boldsymbol{E}_1^v, \boldsymbol{E}_1^w) \\ \text{s.t. } \boldsymbol{X}^v\boldsymbol{W}^v = (\boldsymbol{P}^v\boldsymbol{H} + \boldsymbol{E}_1^v)\boldsymbol{W}^v, \boldsymbol{P}^v\boldsymbol{H} + \boldsymbol{E}_1^v = (\boldsymbol{P}^v\boldsymbol{H} + \boldsymbol{E}_1^v)\boldsymbol{Z}^v + \boldsymbol{E}_2^v, \\ \boldsymbol{H} = \boldsymbol{H}\boldsymbol{Z} + \boldsymbol{E}_H, (\boldsymbol{P}^v)^T\boldsymbol{P}^v = \boldsymbol{I}, \mathcal{Z} = \Phi(\boldsymbol{Z}^1, \boldsymbol{Z}^2, \cdots, \boldsymbol{Z}^V, \boldsymbol{Z})\end{aligned} \tag{6}$$

where $\Upsilon = \{\boldsymbol{H}, \boldsymbol{P}^v, \boldsymbol{E}_1^v, \boldsymbol{Z}, \boldsymbol{E}_H, \boldsymbol{Z}^v, \boldsymbol{E}_2^v\}$ is unknown variable set, λ_1 and λ_2 are hyperparameters that control the contribution of different regularization terms in the objective function. In Eq.(6), the shared latent representation $\boldsymbol{H}$ not only facilitates the joint optimization and collaborative interaction between view recovery and subspace learning but also serves as a semantic anchor that aligns the reconstructed views with corresponding subspace representations in capturing cross-view complementarity and consistency. The information flow propagation enables coherent cross-view semantic correlation exploration, promoting the clustering quality for complex IMVC task.

3.4 Optimization

The objective function in Eq.(6) with multiple variables is difficult to solve directly, the Alternating Direction Method of Multipliers (ADMM) (Lin et al., 2011) is utilized to iteratively optimize each variable. To make the objective function separable, auxiliary variables $\mathcal{J}$ and $\boldsymbol{X}_c^v$ are introduced, and then the augmented Lagrange function is defined:

$$\begin{aligned}\mathcal{L}\big(\Upsilon, \boldsymbol{X}_c^v, \mathcal{J}; \boldsymbol{Y}_1^v, \boldsymbol{Y}_2^v, \boldsymbol{Y}_3^v, \boldsymbol{Y}_4, \mathcal{Y}, \mu\big) = \|\mathcal{J}\|_{\circledast} + \lambda_1(\|\boldsymbol{E}_H\|_{2,1} + \sum_{v=1}^{V}\|\boldsymbol{E}_2^v\|_{2,1}) \\ + \lambda_2 \sum_{w=1;v\neq w}^{V} \text{HSIC}(\boldsymbol{E}_1^v, \boldsymbol{E}_1^w) + \sum_{v=1}^{V}\phi\big(\boldsymbol{Y}_1^v, \boldsymbol{X}_c^v - \boldsymbol{P}^v\boldsymbol{H} - \boldsymbol{E}_1^v\big) \\ + \sum_{v=1}^{V}\phi\big(\boldsymbol{Y}_2^v, \boldsymbol{X}^v\boldsymbol{W}^v - \boldsymbol{X}_c^v\boldsymbol{W}^v\big) + \sum_{v=1}^{V}\phi\big(\boldsymbol{Y}_3^v, \boldsymbol{X}_c^v - \boldsymbol{X}_c^v\boldsymbol{Z}^v - \boldsymbol{E}_2^v\big) \\ + \phi\left(\boldsymbol{Y}_4, \boldsymbol{H} - \boldsymbol{H}\boldsymbol{Z} - \boldsymbol{E}_H\right) + \phi\left(\mathcal{Y}, \mathcal{Z} - \mathcal{J}\right)\end{aligned} \tag{7}$$

where $\{\boldsymbol{Y}_1^v, \boldsymbol{Y}_2^v, \boldsymbol{Y}_3^v, \boldsymbol{Y}_4, \mathcal{Y}\}$ are Lagrangian multipliers, μ is a penalty factor, $\phi(\boldsymbol{A}, \boldsymbol{B}) = \frac{\mu}{2}\|\boldsymbol{B}\|_F^2 + \langle\boldsymbol{A}, \boldsymbol{B}\rangle$. Then, each variable is alternately updated as follows:

Update $\boldsymbol{P}^v$: Fixing other variables except $\boldsymbol{P}^v$, the subproblem in Eq.(7) w.r.t. $\boldsymbol{P}^v$ is as follows:

$$\min_{(\boldsymbol{P}^v)^T\boldsymbol{P}^v=\boldsymbol{I}} \phi(\boldsymbol{Y}_1^v, \boldsymbol{X}_c^v - \boldsymbol{P}^v\boldsymbol{H} - \boldsymbol{E}_1^v) \Longleftrightarrow \min_{(\boldsymbol{P}^v)^T\boldsymbol{P}^v=\boldsymbol{I}} Tr((\boldsymbol{P}^v)^T(\boldsymbol{X}_c^v - \boldsymbol{E}_1^v + \boldsymbol{Y}_1^v/\mu)\boldsymbol{H}^T) \tag{8}$$

The optimal closed-form solution could be obtained (Wang et al., 2019), i.e., $\boldsymbol{P}^v = \boldsymbol{U}\boldsymbol{V}^T$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are the left and right singular vector of $(\boldsymbol{X}_c^v - \boldsymbol{E}_1^v + \boldsymbol{Y}_1^v/\mu)\boldsymbol{H}^T$.

Update $\boldsymbol{X}_c^v$: Fixing other variables except $\boldsymbol{X}_c^v$, the subproblem in Eq.(7) w.r.t. variable $\boldsymbol{X}_c^v$ is as follows:

$$\min_{\boldsymbol{X}_c^v} \phi(\boldsymbol{Y}_1^v, \boldsymbol{X}_c^v - \boldsymbol{P}^v\boldsymbol{H} - \boldsymbol{E}_1^v) + \phi(\boldsymbol{Y}_2^v, \boldsymbol{X}^v\boldsymbol{W}^v - \boldsymbol{X}_c^v\boldsymbol{W}^v) + \phi(\boldsymbol{Y}_3^v, \boldsymbol{X}_c^v - \boldsymbol{X}_c^v\boldsymbol{Z}^v - \boldsymbol{E}_2^v) \tag{9}$$

Taking the derivative w.r.t. $\boldsymbol{X}_c^v$ and setting it to zero, the optimal solution for variable $\boldsymbol{X}_c^v$ is as follows:

$$\boldsymbol{X}_c^v = \big(\boldsymbol{A}_1 + \boldsymbol{A}_2(\boldsymbol{W}^v)^T + \boldsymbol{A}_3(\boldsymbol{I} - \boldsymbol{Z}^v)^T\big)\big(\boldsymbol{I} + \boldsymbol{W}^v(\boldsymbol{W}^v)^T + (\boldsymbol{I} - \boldsymbol{Z}^v)(\boldsymbol{I} - \boldsymbol{Z}^v)^T\big)^{-1} \tag{10}$$

where $\boldsymbol{A}_1 = \boldsymbol{P}^v\boldsymbol{H} + \boldsymbol{E}_1^v - \boldsymbol{Y}_1^v/\mu$, $\boldsymbol{A}_2 = \boldsymbol{X}^v\boldsymbol{W}^v + \boldsymbol{Y}_2^v/\mu$, and $\boldsymbol{A}_3 = \boldsymbol{E}_2^v - \boldsymbol{Y}_3^v/\mu$.

Update $\boldsymbol{H}$: Fixing other variables except $\boldsymbol{H}$, the subproblem in Eq.(7) w.r.t. variable $\boldsymbol{H}$ is as follows:

$$\min_{\boldsymbol{H}} \sum_{v=1}^{V} \phi(\boldsymbol{Y}_1^v, \boldsymbol{X}_c^v - \boldsymbol{P}^v\boldsymbol{H} - \boldsymbol{E}_1^v) + \phi(\boldsymbol{Y}_4, \boldsymbol{H} - \boldsymbol{H}\boldsymbol{Z} - \boldsymbol{E}_H) \tag{11}$$

Taking the derivative w.r.t. variable $\boldsymbol{H}$ and setting it to zero, the optimal solution could be obtained:

$$\boldsymbol{A}\boldsymbol{H} + \mu\boldsymbol{H}\boldsymbol{B} = \boldsymbol{C} \tag{12}$$

where $\boldsymbol{A} = \sum_{v=1}^{V} \mu(\boldsymbol{P}^v)^T\boldsymbol{P}^v$, $\boldsymbol{B} = \mu(\boldsymbol{I} - \boldsymbol{Z} - \boldsymbol{Z}^T + \boldsymbol{Z}\boldsymbol{Z}^T)$, $\boldsymbol{C} = \mu\left(\boldsymbol{E}_H - \boldsymbol{Y}_4/\mu\right)\left(\boldsymbol{I} - \boldsymbol{Z}^T\right) + \sum_{v=1}^{V} \mu(\boldsymbol{P}^v)^T\left(\boldsymbol{X}_c^v - \boldsymbol{E}_1^v + \boldsymbol{Y}_1^v/\mu\right)$. It is a Sylvester equation in Eq.(12) (Bartels & Stewart, 1972). The matrix $\boldsymbol{A}$ is relaxed into $\hat{\boldsymbol{A}} = \boldsymbol{A} + \epsilon\boldsymbol{I}$ with strictly positive definiteness for the solution stability (ϵ is a small positive scalar).

Update $\boldsymbol{Z}^v$: Fixing other variables except $\boldsymbol{Z}^v$, the subproblem in Eq.(7) w.r.t variable $\boldsymbol{Z}^v$ is as follows:

$$\min_{\boldsymbol{Z}^v} \phi(\boldsymbol{Y}_3^v, \boldsymbol{X}_c^v - \boldsymbol{X}_c^v\boldsymbol{Z}^v - \boldsymbol{E}_2^v) + \phi(\boldsymbol{Y}^v, \boldsymbol{Z}^v - \boldsymbol{J}^v) \tag{13}$$

where $\boldsymbol{Y}^v = \Phi_v^{-1}(\mathcal{Y})$, $\boldsymbol{Z}^v = \Phi_v^{-1}(\mathcal{Z})$ and $\boldsymbol{J}^v = \Phi_v^{-1}(\mathcal{J})$. Taking the derivative w.r.t. $\boldsymbol{Z}^v$ and setting it to zero, the optimal solution for variable $\boldsymbol{Z}^v$ could be obtained:

$$\boldsymbol{Z}^v = \left((\boldsymbol{X}_c^v)^T\boldsymbol{X}_c^v + \boldsymbol{I}\right)^{-1}\left((\boldsymbol{X}_c^v)^T\boldsymbol{X}_c^{(v)} - (\boldsymbol{X}_c^v)^T\boldsymbol{E}_2^v + (\boldsymbol{X}_c^v)^T\boldsymbol{Y}_3^v/\mu + \boldsymbol{J}^v - \boldsymbol{Y}^v/\mu\right) \tag{14}$$

Update $\boldsymbol{Z}$: Fixing other variables except $\boldsymbol{Z}$, the subproblem in Eq.(7) w.r.t. variable $\boldsymbol{Z}$ is as follows:

$$\min_{\boldsymbol{Z}} \phi(\boldsymbol{Y}_4, \boldsymbol{H} - \boldsymbol{H}\boldsymbol{Z} - \boldsymbol{E}_H) + \phi(\boldsymbol{Y}, \boldsymbol{Z} - \boldsymbol{J}) \tag{15}$$

where $\boldsymbol{Y} = \Phi_{V+1}^{-1}(\mathcal{Y})$, $\boldsymbol{Z} = \Phi_{V+1}^{-1}(\mathcal{Z})$ and $\boldsymbol{J} = \Phi_{V+1}^{-1}(\mathcal{J})$. Taking the derivative w.r.t. $\boldsymbol{Z}$ and setting it to zero, the optimal solution for variable $\boldsymbol{Z}$ could be obtained:

$$\boldsymbol{Z} = (\boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{I})^{-1}\left(\boldsymbol{H}^T(\boldsymbol{H} - \boldsymbol{E}_H + \boldsymbol{Y}_4/\mu) + \boldsymbol{J} - \boldsymbol{Y}/\mu\right) \tag{16}$$

Update $\boldsymbol{E}_1^v$: Fixing other variables except $\boldsymbol{E}_1^v$, the subproblem in Eq.(7) w.r.t. variable $\boldsymbol{E}_1^v$ is as follows:

$$\min_{\boldsymbol{E}_1^v} \phi(\boldsymbol{Y}_1^v, \boldsymbol{X}_c^v - \boldsymbol{P}^v\boldsymbol{H} - \boldsymbol{E}_1^v) + \lambda_2 \sum_{w=1;w\neq v}^{V} \text{HSIC}(\boldsymbol{E}_1^v, \boldsymbol{E}_1^w) \tag{17}$$

Here, the inner product kernel (i.e., $\boldsymbol{K} = (\boldsymbol{E}_1^v)^T\boldsymbol{E}_1^v$) is utilized for HSIC term. Taking the derivative w.r.t. $\boldsymbol{E}_1^v$ and setting it to zero, the optimal solution for varaible $\boldsymbol{E}_1^v$ could be obtained:

$$\boldsymbol{E}_1^v = \left(\boldsymbol{X}_c^v - \boldsymbol{P}^v\boldsymbol{H} + \boldsymbol{Y}_1^v/\mu\right)\left(\boldsymbol{I} + \frac{2\lambda_2}{\mu(n-1)^2} \sum_{w=1;w\neq v}^{V} \tilde{\boldsymbol{H}}\boldsymbol{K}^w\tilde{\boldsymbol{H}}\right)^{-1} \tag{18}$$

Update $\boldsymbol{E}_2^v$: Fixing other variables except $\boldsymbol{E}_2^v$, the subproblem in Eq.(7) w.r.t. variable $\boldsymbol{E}_2^v$ is as follows:

$$\min_{\boldsymbol{E}_2^v} \lambda_1\|\boldsymbol{E}_2^v\|_{2,1} + \phi(\boldsymbol{Y}_3^v, \boldsymbol{X}_c^v - \boldsymbol{X}_c^v\boldsymbol{Z}^v - \boldsymbol{E}_2^v) \Longleftrightarrow \min_{\boldsymbol{E}_2^v} \lambda_1/\mu\|\boldsymbol{E}_2^v\|_{2,1} + 1/2\|\boldsymbol{E}_2^v - \boldsymbol{L}^v\|_F^2 \tag{19}$$

where $\boldsymbol{L}^v = \boldsymbol{X}_c^v - \boldsymbol{X}_c^v\boldsymbol{Z}^v + \boldsymbol{Y}_3^v/\mu$. Its solution can be obtained by $l_{2,1}$ minimization thresholding operator column by column (Liu et al., 2010), i.e.,

$$\boldsymbol{E}_2^v(:,j) = \left(1 - \frac{\lambda_1}{\mu\left\|\boldsymbol{L}^v(:,j)\right\|_2}\right)^+ \boldsymbol{L}^v(:,j) \tag{20}$$

where $(x)^+ = \max(x, 0)$.

Update $\boldsymbol{E}_H$: Fixing other variables except $\boldsymbol{E}_H$, the subproblem in Eq.(7) w.r.t. variable $\boldsymbol{E}_H$ is as follows:

$$\min_{\boldsymbol{E}_H} \lambda_1\|\boldsymbol{E}_H\|_{2,1} + \phi(\boldsymbol{Y}_4, \boldsymbol{H} - \boldsymbol{H}\boldsymbol{Z} + \boldsymbol{Y}_4/\mu) \tag{21}$$

Similar to $\boldsymbol{E}_2^v$, the solution of $\boldsymbol{E}_H$ could be solved by $l_{2,1}$ minimization thresholding operator column by column, i.e.,

$$\boldsymbol{E}_H(:,j) = \left(1 - \frac{\lambda_1}{\mu \left\|\boldsymbol{M}(:,j)\right\|_2}\right)^+ \boldsymbol{M}(:,j) \tag{22}$$

where $\boldsymbol{M} = \boldsymbol{H} - \boldsymbol{H}\boldsymbol{Z} + \boldsymbol{Y}_4/\mu$.

Update $\mathcal{J}$: Fixing other variables except $\mathcal{J}$, the subproblem in Eq.(7) w.r.t. variable $\mathcal{J}$ is as follows:

$$\min_{\mathcal{J}} \|\mathcal{J}\|_{\circledast} + \phi(\mathcal{Y}, \mathcal{Z} - \mathcal{J}) \Longleftrightarrow \min_{\mathcal{J}} 1/\mu \|\mathcal{J}\|_{\circledast} + 1/2 \|\mathcal{Z} + \mathcal{Y}/\mu - \mathcal{J}\|_F^2 \tag{23}$$

It is a classical tensor nuclear norm minimization problem, where the closed-form solution could be obtained via tensor singular value thresholding (t-SVT) (Zhang et al., 2014).

Update multipliers and penalty parameter: The $\{\boldsymbol{Y}_1^v, \boldsymbol{Y}_2^v, \boldsymbol{Y}_3^v\}_{v=1}^V$, $\boldsymbol{Y}_4$, $\mathcal{Y}$, μ are updated as follows:

$$\begin{cases} \boldsymbol{Y}_1^v := \boldsymbol{Y}_1^v + \mu(\boldsymbol{X}_c^v - \boldsymbol{P}^v\boldsymbol{H} - \boldsymbol{E}_1^v), \\ \boldsymbol{Y}_2^v := \boldsymbol{Y}_2^v + \mu(\boldsymbol{X}^v\boldsymbol{W}^v - \boldsymbol{X}_c^v\boldsymbol{W}^v), \\ \boldsymbol{Y}_3^v := \boldsymbol{Y}_3^v + \mu(\boldsymbol{X}_c^v - \boldsymbol{X}_c^v\boldsymbol{Z}^v - \boldsymbol{E}_2^v), \\ \boldsymbol{Y}_4 := \boldsymbol{Y}_4 + \mu(\boldsymbol{H} - \boldsymbol{H}\boldsymbol{Z} - \boldsymbol{E}_H), \\ \mathcal{Y} := \mathcal{Y} + \mu(\mathcal{Z} - \mathcal{J}), \\ \mu := \min(\rho\mu, \mu_{max}) \end{cases} \tag{24}$$

where $\rho > 1$ is set to accelerate the convergence. With all initial variables, each variable is iteratively updated until stop conditions are satisfied. After obtaining comprehensive subspace representations $(\boldsymbol{Z}, \{\boldsymbol{Z}^v\}_{v=1}^V)$, a powerful affinity $\boldsymbol{S}$ is constructed for spectral clustering, i.e., $\boldsymbol{S} = (|\boldsymbol{Z}| + |\boldsymbol{Z}^T| + \sum_{v=1}^V |\boldsymbol{Z}^v| + \sum_{v=1}^V |(\boldsymbol{Z}^v)^T|)/(V+1)$.

4 Experiments

4.1 Experimental Setting

Datasets: To verify the effectiveness of ARSL-IMVC method, seven benchmark datasets are utilized, including BBCSport (Zhang et al., 2024), HW (Asuncion et al., 2007), BDGP (Cai et al., 2012), Yale (Guo et al., 2025), NGs (Hussain et al., 2010), 100leaves (Mallah et al., 2013), and Scene-15 (Li & Perona, 2005). ***Competitors:*** Several representative methods are selected as competitors, including **BSV** (Ng et al., 2001), **Concat** (Wen et al., 2023b), **IMSC-AGL** (Wen et al., 2020a), **DAIMC** (Hu & Chen, 2018), **UEAF** (Wen et al., 2019), **HCP-IMSC** (Li et al., 2022), **HCLS-CGL** (Wen et al., 2023a), **BWIC-TIMC** (Yao et al., 2025), **RMoGL** (Guo et al., 2025). ***Incomplete Data Construction:*** The incomplete multi-view data is constructed by randomly removing samples from each view, with missing rates $p \in \{0.1, 0.3, 0.5\}$ on the BBCSport, HW, BDGP datasets, missing rates $p \in \{0.1 : 0.1 : 0.8\}$ on the Yale, NGs, 100leaves, and Scene-15 datasets. Following the experimental setting in (Liu et al., 2024b), each sample appears in at least one view. ***Parameters Settings:*** All compared methods are implemented with their public source codes and recommended parameter settings. For ARSL-IMVC, both λ_1 and λ_2 are searched within the range of $\{1, 10, 20, 30, 40, 50\}$, and k is selected between 10 and 20. ***Clustering Metrics:*** To make comprehensive comparison, all methods are repeated ten times and the average values of Accuracy (ACC), Normalized Mutual Information (NMI), and Purity metrics are reported.

4.2 Experimental Results and Analysis

Table 2 presents quantitative clustering results on the BBCSport, HW, and BDGP datasets, where the best and sub-optimal results are marked in bold and underline. Figure 2 illustrates the experimental results with broader missing rates on the Yale, NGs, 100leaves and Scene-15 datasets. The the following conclusions can be concluded:

- The proposed ARSL-IMVC method consistently outperforms other IMVC existing methods on most cases. And, the ARSL-IMVC achieves significant performance improvements. For exapmle, in terms of ACC, it improves around 4.60%, 8.31%, and 5.41% over sub-optimal methods respectively on BBCSport, HW, and BDGP datasets, when missing ratio is 0.1.

Table 2: Clustering results of all methods on the BBCSport, HW, and BDGP datasets.

Dataset	p	Metrics	BSV	Concat	DAIMC	UEAF	IMSC-AGL	HCLS-CGL	HCP-IMSC	BWIC-TIMC	RMoGL	**Ours**
BBCSport	0.1	ACC	35.85	58.07	75.39	71.32	84.30	72.00	91.91	90.75	89.19	**96.51**
		NMI	1.98	34.96	60.76	63.27	72.63	68.90	79.84	82.10	83.18	**89.77**
		Purity	36.21	59.52	78.18	80.70	85.75	75.55	91.91	90.34	88.79	**96.51**
	0.3	ACC	33.82	46.27	75.31	77.39	88.27	79.96	89.15	78.94	78.49	**94.85**
		NMI	1.82	17.69	58.29	58.19	74.34	71.33	75.47	70.12	68.86	**84.95**
		Purity	36.21	47.81	77.46	79.78	88.62	83.45	89.15	81.25	81.80	**94.85**
	0.5	ACC	30.51	45.44	57.74	67.22	81.99	78.49	86.05	77.53	76.47	**88.97**
		NMI	2.32	15.81	41.25	52.05	66.79	67.13	**72.37**	65.84	61.76	71.32
		Purity	36.21	46.69	66.62	73.29	82.52	82.17	86.05	79.48	79.60	**88.97**
HW	0.1	ACC	44.10	65.08	76.65	53.79	88.59	79.81	79.80	81.85	76.73	**96.90**
		NMI	52.11	63.83	71.59	50.04	87.21	81.79	75.73	82.87	74.61	**92.77**
		Purity	44.20	68.99	78.99	54.01	90.67	81.97	80.05	81.85	76.73	**96.90**
	0.3	ACC	38.50	57.05	59.18	44.48	84.28	81.45	75.35	74.75	64.19	**92.46**
		NMI	43.11	52.25	51.45	40.66	78.31	81.04	69.11	72.63	62.25	**84.16**
		Purity	38.90	58.97	60.16	45.35	85.64	84.00	76.50	74.75	65.03	**92.46**
	0.5	ACC	31.75	44.86	62.59	37.07	79.43	81.75	70.80	60.13	59.20	**89.03**
		NMI	33.78	39.15	51.01	30.85	73.36	**81.39**	60.47	59.70	54.74	77.97
		Purity	32.00	47.15	63.28	37.27	81.34	84.00	71.30	62.64	59.35	**89.03**
BDGP	0.1	ACC	40.28	45.02	45.72	50.66	41.67	23.68	21.08	29.88	45.94	**56.07**
		NMI	25.22	22.13	22.91	28.03	19.90	3.37	25.26	7.23	23.48	**35.22**
		Purity	46.08	46.58	47.75	52.42	44.57	23.68	19.52	31.04	46.66	**56.07**
	0.3	ACC	39.22	39.20	41.85	46.82	38.97	23.84	23.93	20.32	42.66	**50.58**
		NMI	23.31	16.55	17.03	22.58	16.93	3.36	29.06	1.30	20.30	**31.59**
		Purity	43.52	40.71	42.98	49.86	41.55	23.84	22.21	20.36	43.60	**52.07**
	0.5	ACC	37.68	34.90	39.74	45.92	37.59	24.28	20.46	25.52	31.68	**49.21**
		NMI	20.99	11.78	20.22	24.14	15.81	3.57	24.84	2.43	6.77	**32.16**
		Purity	40.60	35.43	41.87	47.78	39.42	24.32	19.00	25.61	32.68	**50.21**

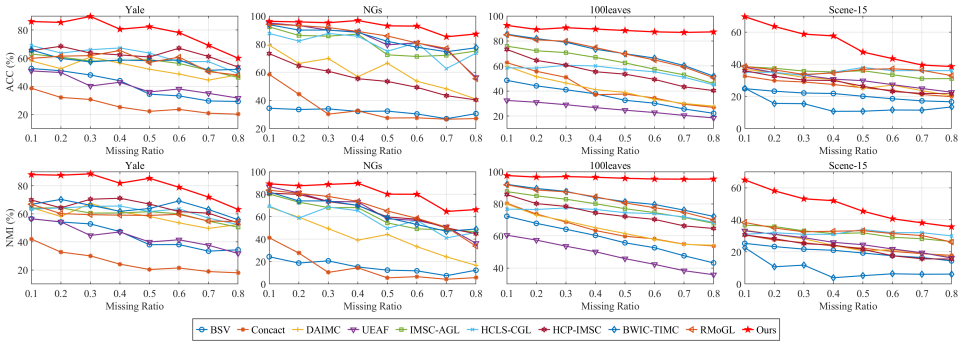
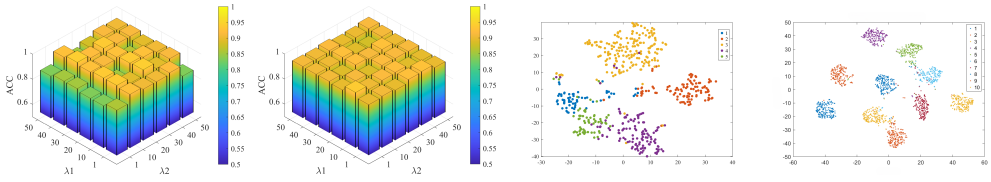


Figure 2: Clustering results of all methods on Yale, NGs, 100leaves and Scene-15 with different missing rates.

- Compared to imputation-free IMVC methods (DAIMC, IMSC-AGL, HCLS-CGL), the proposed ARSL-IMVC usually achieves superior clustering performance, illustrating the effectiveness of proposed collaborative view recovery strategy. And, it highlights that the reshaping cross-view consistency and diversity in feature-level indeed facilitates credible and strong view recovery.
- The proposed ARSL-IMVC method also outperforms the imputation-based methods (i.e., UEAF, HCP-IMSC, BWIC-TIMC, and RMoGL), owing to its learned latent representation that simultaneously serves as a semantic foundation for both view recovery and subspace learning, thereby enabling their alignment in capturing cross-view consistency and complementarity. This explicit information flow transmission promotes deep interaction between CVR and TSL modules, improving the ability to accurately recover missing view and clustering semantic discriminability.
- As illustrated in Figure 2, most IMVC methods exhibit significant performance degradation as the missing rate increases, while ARSL-IMVC maintains higher stability. It fully demonstrates the superiority of proposed ARSL-IMVC in complex IMVC tasks.

Table 3: The experimental results of ARSL-IMVC and its ablation variant with the 0.1 missing ratio.

Datasets	BBCSport		HW		Yale		NGs		100leaves	
Metrics	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI
ARSL-IMVC-1	84.03	71.52	70.15	61.77	76.55	78.88	89.96	76.65	78.61	91.09
ARSL-IMVC	**96.51**	**89.77**	**96.90**	**92.77**	**86.06**	**87.96**	**96.20**	**89.27**	**89.24**	**96.65**



(a) Yale (b) 100leaves (c) BBCSport (d) HW

Figure 3: Parameter sensitivity analysis of the λ_1 and λ_2 on Yale and 100leaves datasets with the missing rate of 0.1 in Figures (3a) and (3b). The t-SNE visualization of clustering representation on BBCSport and HW with the missing rate of 0.1 in Figures (3c) and (3d).

4.3 Ablation Study

To verify effectiveness of aligning the view recovery and subspace learning by latent representation $\boldsymbol{H}$, one ablation variant ARSL-IMVC-1 is designed by removing the subspace learning on $\boldsymbol{H}$.

As shown in the Table 3, ARSL-IMVC consistently outperforms ablation variant ARSL-IMVC-1, achieving a performance improvement of 12.48%, 26.75%, 9.51%, 6.24%, and 10.63% on BBCSport, HW, Yale, NGs, and 100leaves datasets regarding ACC metric. The result indicates that aligning view recovery and subspace learning in complex correlation exploration promotes the semantic coherence between them, facilitating the fine-grained and clear clustering structure modeling.

4.4 Parameter Sensitivity and Visualization Analysis

To evaluate the sensitivity of ARSL-IMVC to parameter, with a missing rate of 0.1, the ACC metric with different λ_1 and λ_2 is shown in Figures (3a) and (3b). It can be observed that the performance of ARSL-IMVC is not significantly influenced by λ_2 when λ_1 is fixed, while it is slightly affected by λ_1 when λ_2 is fixed. Although, the ARSL-IMVC is relatively robust to parameter varying within a reasonable range.

To intuitively illustrate the clustering performance of ARSL-IMVC, the spectral embeddings on BBCSport and HW datasets are visualized by t-SNE (Maaten & Hinton, 2008) with a missing rate of 0.1. As shown in Figures (3c) and (3d), the spectral embedding obtained by ARSL-IMVC exhibits relatively clear clustering structure and samples from diverse clusters are obviously separated, verifying its discriminability in clustering semantics exploration.

4.5 Scalability on Large-scale Dataset

To further validate the effectiveness and scalability of the proposed ARSL-IMVC on large-scale multi-view data, the Handwritten Digits (HDigit) dataset with 10000 samples is utilized, where two views are collected from various resources: MNIST and USPS. The experimental results of partial representative methods in 0.1 missing ratio are shown in Table 4. Compared with the suboptimal HCLS-IMSC, improvements of approximately 0.7% in both ACC and Purity and approximately 1.7% in NMI are achieved by the proposed method. Compared to other baselines such as UEAF and IMSC-AGL, the performance gains reach $10\% \sim 25\%$. These results clearly demonstrate that ARSL-IMVC can effectively handle the representation learning challenges arising from substantial increases in sample size and feature dimensionality.

4.6 Running Time Comparison and Convergence Analysis

Table 4: Clustering results of some methods on **HDigit** datasets with the missing rate of 0.1.

Methods	DAIMC	UEAF	IMSC-AGL	HCLS-IMSC	HCP-IMSC	**Ours**
ACC	67.58	85.38	76.32	98.29	89.56	**99.00**
NMI	64.25	73.56	77.40	95.30	89.52	**96.97**
Purity	69.61	85.38	77.89	98.29	88.40	**99.00**

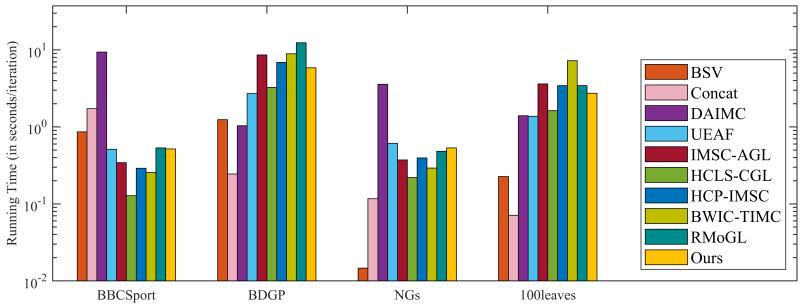


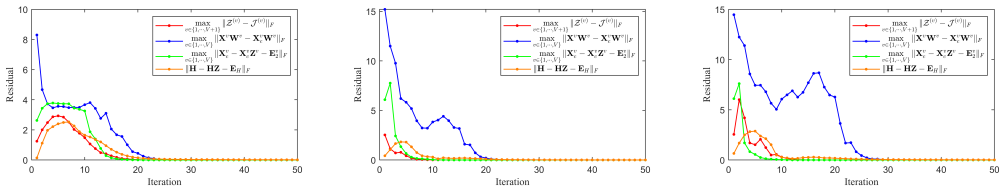
Figure 4: The running time comparisons of all methods on the BBCSport, BDGP, NGs and 100leaves datasets with 0.1 missing rate.

Figure 4 reports the running time of all compared methods on the four benchmark datasets. As shown in Figure 4, the proposed ARSL-IMVC method generally requires comparable computational costs to other IMVC methods on the four datasets. Considering the superior clustering performance of proposed method, ARSL-IMVC demonstrates good balance between computational efficiency and clustering performance.

To empirically shown the stability of the proposed algorithm, the convergence analysis results are verified on the Yale, BBCSport, and NGs datasets with the missing rate of 0.1, where the iterative residual errors are visualized. As shown in Figures (5a), (5b) and (5c), the residual curves of variable updating show that proposed optimization algorithm can reach a local minimum within a finite number of iterations, verifying the fast convergence and numerical stability of ARSL-IMVC.

5 Conclusion

In this study, a novel ARSL-IMVC method was proposed for incomplete multi-view clustering, which centers on a latent representation to achieve unified alignment between view recovery and tensor subspace learning in complex cross-view correlation exploration. The latent representation served as both view reconstruction basis and global semantics carrier, maintaining and aligning the cross-view consistency. And, global view-shared and local view-specific subspace representations were organized into a low-rank tensor, exploring the cross-view complementarity and multi-level structural correlations. Experimental results demonstrated that ARSL-IMVC consistently achieves superior clustering performance under various missing rates.



(a) Yale (b) BBCSport (c) NGs

Figure 5: The convergence curves on Yale, BBCSport and NGs datsets with the missing rate of 0.1.

ETHICS STATEMENT

Our research adheres strictly to ICLR ethical guidelines, emphasizing responsible AI development to maximize societal benefits while mitigating potential harms. The proposed ARSL-IMVC method enhances incomplete multi-view clustering, offering advancements in data analysis that could benefit fields such as healthcare, environmental monitoring, and education by improving resilience to missing data. Our research is dedicated to the benefit of society and human well-being, and since it does not involve human subjects, no ethical approval is required. All datasets used in our experiments are publicly available and comply with their respective licenses. The methodology and results are presented with full transparency to support reproducibility, including source code provided in the supplementary materials, which will be open-sourced following the review process.

REPRODUCIBILITY STATEMENT

To ensure the reproducibility of our work, the complete source code of the ARSL-IMVC will be provided in the supplementary materials. The public datasets used in the experiments are all cited in the paper and the construction of incomplete datasets can be carried out according to the specific methods mentioned in Section 4.1 of the main text.

REFERENCES

Arthur Asuncion, David Newman, et al. Uci machine learning repository, 2007.

Shunshun Bai, Qinghai Zheng, Xiaojin Ren, and Jihua Zhu. Graph-guided imputation-free incomplete multi-view clustering. *Expert Systems with Applications*, 258:125165, 2024.

Richard H Bartels and George W Stewart. Algorithm 432 [c2]: Solution of the matrix equation ax+ xb= c [f4]. *Communications of the ACM*, 15(9):820–826, 1972.

Xiao Cai, Hua Wang, Heng Huang, and Chris Ding. Joint stage recognition and anatomical annotation of drosophila gene expression patterns. *Bioinformatics*, 28(12):16–24, 2012.

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*, pp. 586–594, 2015.

Guoqing Chao, Shiliang Sun, and Jinbo Bi. A survey on multiview clustering. *IEEE Transactions on Artificial Intelligence*, 2(2):146–168, 2021.

Jin Chen, Huafu Xu, Jingjing Xue, Quanxue Gao, Cheng Deng, and Ziyu Lv. Incomplete multi-view clustering based on hypergraph. *Information Fusion*, 117:102804, 2025.

Man-Sheng Chen, Jia-Qi Lin, Xiang-Long Li, Bao-Yu Liu, Chang-Dong Wang, Dong Huang, and Jian-Huang Lai. Representation learning in multi-view clustering: A literature review. *Data Science and Engineering*, 7(3):225–241, 2022.

Yongyong Chen, Xiaojia Zhao, Zheng Zhang, Youfa Liu, Jingyong Su, and Yicong Zhou. Tensor learning meets dynamic anchor learning: From complete to incomplete multiview clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 35(11):15332–15345, 2024.

Hongchang Gao, Feiping Nie, Xuelong Li, and Heng Huang. Multi-view subspace clustering. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4238–4246, 2015.

Arthur Gretton, Kenji Fukumizu, Choon Teo, Le Song, Bernhard Schölkopf, and Alex Smola. A kernel statistical test of independence. In *Advances in Neural Information Processing Systems*, pp. 585–592, 2007.

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

Wei Guo, Hangjun Che, Man-Fai Leung, Long Jin, and Shiping Wen. Robust mixed-order graph learning for incomplete multi-view clustering. *Information Fusion*, 115:102776, 2025.

Menglei Hu and Songcan Chen. Doubly aligned incomplete multi-view clustering. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 2262–2268, 2018.

Syed Fawad Hussain, Gilles Bisson, and Clément Grimal. An improved co-similarity measure for document clustering. In *International Conference on Machine Learning and Applications*, pp. 190–197, 2010.

Bingbing Jiang, Chenglong Zhang, Xinyan Liang, Peng Zhou, Jie Yang, Xingyu Wu, Junyi Guan, Weiping Ding, and Weiguo Sheng. Collaborative similarity fusion and consistency recovery for incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 17617–17625, 2025.

Fei F. Li and Pietro Perona. A bayesian hierarchical model for learning natural scene categories. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 524–531, 2005.

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.

ShaoYuan Li, Yuan Jiang, and ZhiHua Zhou. Partial multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 1968–1974, 2014.

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.

Zhouchen Lin, Risheng Liu, and Zhixun Su. Linearized alternating direction method with adaptive penalty for low-rank representation. In *Advances in Neural Information Processing Systems*, pp. 612–620, 2011.

Chengliang Liu, Jie Wen, Zhihao Wu, Xiaoling Luo, Chao Huang, and Yong Xu. Information recovery-driven deep incomplete multiview clustering network. *IEEE Transactions on Neural Networks and Learning Systems*, 35(11):15442–15452, 2024a.

Guangcan Liu, Zhouchen Lin, and Yong Yu. Robust subspace segmentation by low-rank representation. In *Proceedings of the International Conference on Machine Learning*, pp. 663–670, 2010.

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*, pp. 14017–14025, 2024b.

Xinwang Liu, Xinzhong Zhu, Miaomiao Li, Lei Wang, En Zhu, Tongliang Liu, Marius Kloft, Dinggang Shen, Jianping Yin, and Wen Gao. Multiple kernel k k-means with incomplete kernels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(5):1191–1204, 2020.

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.

Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9:2579–2605, 2008.

Charles Mallah, James Cope, James Orwell, et al. Plant leaf classification using probabilistic integration of shape, texture and margin features. *Signal Processing, Pattern Recognition and Applications*, 5(1):45–54, 2013.

Andrew Ng, Michael Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems*, pp. 849–856, 2001.

Yalan Qin, Guorui Feng, and Xinpeng Zhang. Scalable one-pass incomplete multi-view clustering by aligning anchors. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 20042–20050, 2025.

Qiangqiang Shen, Xuanqi Zhang, Shuqin Wang, Yuanman Li, Yongsheng Liang, and Yongyong Chen. Dual completion learning for incomplete multi-view clustering. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 9(1):455–467, 2025.

Long Shi, Lei Cao, Jun Wang, and Badong Chen. Enhanced latent multi-view subspace clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(12):12480–12495, 2024.

Qianqian Wang, Huanhuan Lian, Gan Sun, Quanxue Gao, and Licheng Jiao. icmsc: Incomplete cross-modal subspace clustering. *IEEE Transactions on Image Processing*, 30:305–317, 2021.

Siwei Wang, Xinwang Liu, En Zhu, Chang Tang, Jiyuan Liu, Jingtao Hu, Jingyuan Xia, and Jianping Yin. Multi-view clustering via late fusion alignment maximization. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 3778–3784, 2019.

Ziyu Wang, Lusi Li, Xin Ning, Wenkai Tan, Yongxin Liu, and Houbing Song. Incomplete multi-view clustering via structure exploration and missing-view inference. *Information Fusion*, 103: 102123, 2024.

Jie Wen, Zheng Zhang, Yong Xu, Bob Zhang, Lunke Fei, and Hong Liu. Unified embedding alignment with missing views inferring for incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 5393–5400, 2019.

Jie Wen, Yong Xu, and Hong Liu. Incomplete multiview spectral clustering with adaptive graph learning. *IEEE Transactions on Cybernetics*, 50(4):1418–1429, 2020a.

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, 2020b.

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 Conference on Computer Vision and Pattern Recognition*, pp. 15712–15721, 2023a.

Jie Wen, Zheng Zhang, Lunke Fei, Bob Zhang, Yong Xu, Zhao Zhang, and Jinxing Li. A survey on incomplete multiview clustering. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(2):1136–1149, 2023b.

Jie Wen, Gehui Xu, Zhanyan Tang, Wei Wang, Lunke Fei, and Yong Xu. Graph regularized and feature aware matrix factorization for robust incomplete multi-view clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(5):3728–3741, 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*, pp. 15952–15960, 2024.

Tingting Wu, Zhendong Li, Zhibin Gu, Jiazheng Yuan, and Songhe Feng. Koala: Kernel coupling and element imputation induced multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 21581–21589, 2025.

Chang Xu, Dacheng Tao, and Chao Xu. A survey on multi-view learning. *arXiv preprint arXiv:1304.5634*, 2013.

Mingze Yao, Huibing Wang, Yawei Chen, and Xianping Fu. Between/within view information completing for tensorial incomplete multi-view clustering. *IEEE Transactions on Multimedia*, 27:1538–1550, 2025.

Chenglong Zhang, Yang Fang, Xinyan Liang, Han Zhang, Peng Zhou, Xingyu Wu, Jie Yang, Bingbing Jiang, and Weiguo Sheng. Efficient multi view unsupervised feature selection with adaptive structure learning and inference. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 5443–5452, 2024.

Zemin Zhang, Gregory Ely, Shuchin Aeron, Ning Hao, and Misha Kilmer. Novel methods for multilinear data completion and de-noising based on tensor-svd. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3842–3849, 2014.

A APPENDIX

The Use of Large Language Models (LLMs): In this research, large language models (LLMs) were employed exclusively to improve the clarity and grammatical accuracy of the manuscript. Their use was limited to refining sentence structure and correcting syntax to enhance readability and professionalism. At no stage did these tools influence the scientific content, methodology, or results. All core ideas and analyses presented are the original work of the authors, with no LLM-generated content contributing to the intellectual substance of the paper.