# Learn Multi-task Anchor: Joint View Imputation and Label Generation for Incomplete Multi-view Clustering

**Xinxin Wang**[1] , **Yongshan Zhang**[2] , **Yicong Zhou**[1*]
[1]Department of Computer and Information Science, University of Macau
[2]School of Computer Science, China University of Geosciences
xinxinwang1024@gmail.com, yszhang.cug@gmail.com, yicongzhou@um.edu.mo

## Abstract

Anchor-based incomplete multi-view clustering methods utilize anchors to uncover clustering structures. However, relying on anchor graphs for producing final indicators is indirect, which can lead to information loss and suboptimal outcomes. Besides, most methods neglect the potential of anchors for imputing missing views. To address these limitations, we propose a Joint View Imputation and Label Generation (JVILG) method. JVILG comprises the Anchor-based tensorized Label Generation (ALG) module for generating clustering labels and the Anchor-based sparse regularized Subspace Correlation (ASC) module for recovering missing views. The ALG module explicitly connects data observations, the fine-grained anchor matrix, and soft label matrices within a reconstruction framework through a membership matrix, while imposing tensor Schatten p-norm regularization on the constructed label tensor to capture spatial correlations among views. Meanwhile, the ASC module directly uses fine-grained anchors to impute missing data in respective views. By integrating the ALG and ASC modules, JVILG enhances synergy between different tasks and mitigates the impact of missing information on clustering. Experimental results on six datasets demonstrate the effectiveness of JVILG compared to both shallow and deep state-of-the art methods. The code is available at https://github.com/W-Xinxin/JVILG.

## 1 Introduction

Multi-view data has attracted significant attention in many real-world applications such as medical diagnosis [Bai *et al.*, 2024] and multimedia recommendation [Yu *et al.*, 2023]. It captures diverse and complementary information about the same object from different sources. However, equipment malfunctions or selective information collection can result in missing partial views, leading to incomplete multi-view data. For example, in imaging examinations like X-ray, CT, and MRI, patients often undergo partial tests based on specific symptoms, resulting in incomplete medical histories. As one of the most effective tools to analyze incomplete multi-view data, incomplete multi-view clustering groups unlabeled samples into semantic clusters by exploring the correlations between samples and views [Wen *et al.*, 2023a].

In recent years, numerous incomplete multi-view clustering methods have been proposed [Wu *et al.*, 2024; Wan *et al.*, 2024b; Lu *et al.*, 2024; Wen *et al.*, 2023b; Hu and Chen, 2018; Wang *et al.*, 2021a]. Among these, anchor-based incomplete multi-view clustering explores and exploits the representational capacity of anchors to reveal underlying clustering structures, achieving great success. For instance, [Wang *et al.*, 2022] utilized view-shared anchors to capture the consistent data distributions among views. [Liu *et al.*, 2022] introduced view-specific anchors to integrate the complementary information among views. [Li *et al.*, 2023a] employed the distribution of observed data to improve anchor learning. [Li *et al.*, 2024] removed hyperparameters to enhance model applicability. However, these approaches still experience several limitations.

On one hand, most anchor-based incomplete multi-view clustering methods rely on constructing additional anchor graphs to capture relationships among samples. Spectral analysis and factorization techniques are then applied to these graphs to generate final clustering indicators [Wang *et al.*, 2020b]. However, the introduction of additional variables or separate processes can lead to information loss and suboptimal performance [Liu *et al.*, 2021b]. On the other hand, how missing views are handled significantly affects model performance. Some methods utilize only observed views [Wen *et al.*, 2021a; Wang *et al.*, 2021b; Wan *et al.*, 2024b], ignoring the semantic information of the missing views. Some approaches fill in missing views using all available observed views [Wen *et al.*, 2019; Wen *et al.*, 2021b; Liu *et al.*, 2021a], which can introduce redundant information and diminish the quality of the imputed views.

To address these issues, we propose a Joint View Imputation and Label Generation (JVILG) method that facilitates high-quality missing view imputation and direct label generation. First, we project the multi-view data from the original feature space into a latent embedding space, where data reconstruction is learned with the help of fine-grained anchors in respective views. These anchor representations in each

*Corresponding author.

view are explored with orthogonality constraints. For label generation, we construct a view-specific membership matrix that correlates the anchor matrix with cluster centroids, establishing explicit connections among fine-grained anchors, cluster centroids, and soft clustering labels. To further capture the diversity and complementary information among cluster indicators, we impose tensor Schatten p-norm regularization on the constructed label tensor. For missing view imputation, we assume that the missing views share latent subspace structures with fine-grained anchors. Therefore, missing views can be recovered using the linear combinations of anchors from the same subspace. The $L_1$-norm is applied to coefficient matrices to preserve only the necessary intrinsic relationships. Additionally, we introduce adaptive weights to balance the contributions from each view. We devise an alternating optimization algorithm to solve the unified objective function. Our method directly outputs soft labels without post-processing. The main contributions are summarized as follows:

- We propose a novel mutli-task framework for missing view imputation and incomplete multi-view clustering. To our knowledge, this is the first approach that directly generates clustering labels from anchors and recover missing views.
- To facilitate direct label generation, we utilize a membership matrix to establish explicit connections among data observations, fine-grained anchors, and cluster partitioning within a reconstruction framework.
- To effectively impute missing views, we use fine-grained anchors to linearly reconstruct the missing views, applying a sparse constraint on the coefficient matrix to avoid spurious connections.
- Comprehensive experimental results on six datasets demonstrate the advantages of our JVILG over both shallow and deep methods for incomplete multi-view clustering.

## 2 Related Work

### 2.1 Anchor-based Incomplete Multi-view Clustering

Anchors have emerged as powerful tools in effectively revealing data distribution in multi-view clustering. Typically, the $K$-means method and random sampling are used to generate these anchors with clustering potential [Kang *et al.*, 2020]. Recently, [Guo and Ye, 2019] selected common instances as anchors to bridge non-overlapping partial views, though this method only addresses two-view cases. To mitigate the occasionality and randomness associated with heuristic anchors, [Chen *et al.*, 2023] proposed dynamic anchor learning to enhance clustering performance. To address the cross-view anchor misalignment problem, [Li *et al.*, 2023b] introduced a predefined anchor-level graph, transforming this issue into a cross-view cluster-level alignment problem. Additionally, [Zhao *et al.*, 2023] integrated single-view local anchors with concatenated-view global anchors to explore diverse and consistent information among views. Other developments include view-independent anchors [Liu *et al.*, 2022], view-shared anchors [Wang *et al.*, 2022], and deep anchors [Cui *et al.*, 2023], all aimed at improving performance. These methods have demonstrated significant success, highlighting the substantial potential of anchors. However, they depend on additional anchor graphs to generate final labels and fail to directly produce labels from anchors. Moreover, these methods utilize anchors solely to construct anchor graphs, ignoring the potential of anchors for imputing missing views.

### 2.2 Label Generation for Multi-view Clustering

To obtain labels for multi-view data, two primary approaches are commonly used. The first involves applying $K$-means on sample representations to derive final labels, though this method is sensitive to initialization. Some studies utilize matrix factorization to implement $K$-means. For example, [Cai *et al.*, 2013] proposed parallelized multi-view $K$-means to directly output final labels, while [Hu and Chen, 2019] adapted this approach for incomplete multi-view data using an indicate matrix strategy. However, the fixed-dimensional centroid matrix may limit model performance due to its restricted representational capacity [Wan *et al.*, 2023]. The second approach relies on spectral analysis, which constructs graphs to capture correlations among samples. [Tang *et al.*, 2022] expanded multi-view spectral clustering to generate discrete labels, while [Wen *et al.*, 2023a] learned a consensus graph for direct label acquisition using $k$-connected components. However, tuning parameter settings to ensure appropriate $k$-connected components remains challenging. Notably, none of these methods can directly learn labels from incomplete multi-view data using anchors, limiting their ability to reduce information loss.

## 3 Proposed Method

This section elaborates our Joint View Imputation and Label Generation (JVILG) approach. We model the JVILG as a reconstruction problem that leverages subspace correlation to strengthen the relationships between missing views and anchors. We present our approach in two aspects and formulate a unified objective function for JVILG.

### 3.1 Anchor-based Tensorized Label Generation (ALG) for Incomplete Multi-view Data

Much redundancy in original data hinders the ability to depict the underlying data structure using all data samples [Chen *et al.*, 2024]. Anchor-based incomplete multi-view clustering methods rely on anchor representations to enhance the model's discriminative ability [Guo and Ye, 2019]. However, existing methods fail to establish explicit connections among the input observations and underlying data structure. The introduction of mediating variables, such as anchor graphs, can lead to information loss and suboptimal outcomes [Chen *et al.*, 2023; Wan *et al.*, 2024a].

To address this issue, we develop a label generation module based on anchor representations. Given an incomplete multi-view dataset $\left\{X^{(v)} \in \mathbb{R}^{d_v \times n}\right\}_{v=1}^{m}$, where $d_v$ denotes the dimensionality of the $v$-th view, $n$ is the total sample size, and $m$ is the number of views, with missing views zero-imputed,

the objective function can be formulated as follows:

$$\min_{\mathbf{P}_v,\mathbf{A}_v,\mathbf{W}_v,\mathbf{F}_v} \sum_{v=1}^{m} \left\|\mathbf{P}_v\mathbf{X}_v - \mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T\right\|_F^2 + \psi \left\|\mathbf{F}_v\right\|_F^2, \quad \text{s.t.} \quad \mathbf{P}_v^T\mathbf{P}_v = \mathbf{I}, \mathbf{A}_v^T\mathbf{A}_v = \mathbf{I}, \mathbf{W}_v^T\mathbf{1} = \mathbf{1}, \mathbf{W}_v \geq 0, \mathbf{F}_v \geq 0, \tag{1}$$

where $\mathbf{P}_v \in \mathbb{R}^{l\times d_v}$ represents projection matrix, and the uncorrelated constraint enhances the separability of the projected data points. $\mathbf{A}_v \in \mathbb{R}^{l\times b}$ is anchor matrix of the $v$-th view, on which an orthogonality constraint is imposed to achieve mutual independence among different anchor representations, enhancing their discriminative power. $\mathbf{W}_v \in \mathbb{R}^{b\times k}$ is a membership matrix that indicates the degree to which fine-grained anchors are related to the cluster centroids, with $\mathbf{A}_v\mathbf{W}_v$ representing the centroids. $\mathbf{F}_v \in \mathbb{R}^{n\times k}$ is cluster indicator matrix. $\psi$ is a nonnegative hyper-parameter that controls the trade-off.

However, Eq. (1) reconstructs different views independently, without considering their inter-view correlations. Inspired by the excellent performance of the tensor Schatten p-norm [Gao *et al.*, 2020] in capturing inter-view complementary information and spatial structure, we apply it to a third-order tensor, where $F_v$ as lateral slices. Thus Eq. (1) is rewritten as follows:

$$\min_{\mathbf{P}_v,\mathbf{A}_v,\mathbf{W}_v,\mathbf{F}_v} \sum_{v=1}^{m} \left\|\mathbf{P}_v\mathbf{X}_v - \mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T\right\|_F^2 + \psi \left\|\boldsymbol{\mathcal{F}}\right\|_{\circledS p}^p, \quad \text{s.t.} \quad \mathbf{P}_v^T\mathbf{P}_v = \mathbf{I}, \mathbf{A}_v^T\mathbf{A}_v = \mathbf{I}, \mathbf{W}_v^T\mathbf{1} = \mathbf{1}, \mathbf{W}_v \geq 0, \mathbf{F}_v \geq 0. \tag{2}$$

Each frontal slice of $\boldsymbol{\mathcal{F}}$ represents the relationships between sample and cluster centroids across different views, Shatten $p$-norm in Eq. (2) ensures that indicators of different views exhibit across-view similarity.

### 3.2 Anchor-based Sparse Regularized Subspace Correlation (ASC) for View Imputation

To reduce the impact of missing views, JVILG employs dynamic data imputation. High-quality anchors not only disclose the underlying data structure but also aid in recovering missing views. Thus, we can infer the missing views using anchors by solving the following subspace correlation learning module.

$$\min_{\mathbf{E}_v,\mathbf{Z}_v} \beta \left\|\mathbf{E}_v - \mathbf{A}_v\mathbf{Z}_v\right\|_F^2 + \left\|\mathbf{Z}_v\right\|_1, \tag{3}$$

where $\mathbf{E}_v \in \mathbb{R}^{l\times n_v}$ represents missing view matrix, and $\mathbf{Z}_v$ is coefficient matrix that captures subspace correlations. $L_1$-norm regularization term ensures that $\mathbf{Z}_v$ is sparse, maintaining only the essential connections. The missing view can be recovered using a linear combination of a small number of anchor points from the same subspace. $\beta$ is a nonnegative hyper-parameter that controls the trade-off. Unlike existing models that uses the whole observed data to recover missing views, Eq. (3) alleviates the impact of potential noisy observations.

### 3.3 Unified Objective Function

Considering the objectives of label generation and missing view imputation simultaneously, we minimize the following objective function:

$$\begin{aligned}\mathcal{J}(\Psi) = \sum_{v=1}^{m} a_v^2 (&\left\|\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v - \mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T\right\|_F^2 \\ &+ \beta \left\|\mathbf{E}_v - \mathbf{A}_v\mathbf{Z}_v\right\|_F^2 + \left\|\mathbf{Z}_v\right\|_1) + \psi \left\|\boldsymbol{\mathcal{F}}\right\|_{\circledS p}^p, \\ \text{s.t.} \quad &\mathbf{P}_v^T\mathbf{P}_v = \mathbf{I}, \mathbf{A}_v^T\mathbf{A}_v = \mathbf{I}, \mathbf{W}_v^T\mathbf{1} = \mathbf{1}, \mathbf{W}_v \geq 0, \\ &\mathbf{F}_v \geq 0, \mathbf{a1} = 1, a_v \geq 0.\end{aligned} \tag{4}$$

where $\Psi = \{a_v, \mathbf{P}_v, \mathbf{A}_v, \mathbf{W}_v, \mathbf{F}_v, \mathbf{E}_v, \mathbf{Z}_v\}$. $\beta$ is a trade-off hyper-parameter. $a_v^2$ is an adaptive weight for the $v$-th view. $\mathbf{G}_v \in \{0,1\}^{n_v\times n}$ is a prior matrix defined as follows, which indicates the missing information of the $v$-th view:

$$\mathbf{G}_v(i,j) = \begin{cases} 1, & \text{if } h_i^{(v)} = j, \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where $h^{(v)} \in \mathbb{R}^{n_v}$ contains the indices of the $n_v$ missing instances of the $v$-th view, corresponding to their original positions. Based on this, $\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v$ represents a complete view, with all views aligned.

### 3.4 Optimization

Eq. (4) is a non-convex optimization problem, and it is difficult to directly solve these variables. Inspired by Alternating Direction Method of Multipliers (ADMM) [Boyd *et al.*, 2011], we devise an alternating optimization algorithm to address this objective function. By introducing an auxiliary variable $\boldsymbol{\mathcal{J}}$ and setting $\boldsymbol{\mathcal{F}} = \boldsymbol{\mathcal{J}}$, the objective function in Eq. (4) can be expressed in the following separable form:

$$\begin{aligned}\mathcal{J}(\Psi) = \sum_{v=1}^{m} a_v^2 (&\left\|\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v - \mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T\right\|_F^2 \\ &+ \beta \left\|\mathbf{E}_v - \mathbf{A}_v\mathbf{Z}_v\right\|_F^2 + \left\|\mathbf{Z}_v\right\|_1) + \psi \left\|\boldsymbol{\mathcal{J}}\right\|_{\circledS p}^p, \\ &+ \frac{\rho}{2}\left\|\boldsymbol{\mathcal{F}} - \boldsymbol{\mathcal{J}}\right\|_F^2 + \langle \boldsymbol{\mathcal{Y}}, \boldsymbol{\mathcal{F}} - \boldsymbol{\mathcal{J}}\rangle \\ \text{s.t.} \quad &\mathbf{P}_v^T\mathbf{P}_v = \mathbf{I}, \mathbf{A}_v^T\mathbf{A}_v = \mathbf{I}, \mathbf{W}_v^T\mathbf{1} = \mathbf{1}, \mathbf{W}_v \geq 0, \\ &\mathbf{F}_v \geq 0, \mathbf{a1} = 1, a_v \geq 0.\end{aligned} \tag{6}$$

where $\boldsymbol{\mathcal{Y}}$ represents the Lagrange multiplier and $\rho$ denotes the penalty factor. The solution to Eq. (6) can be obtained by solving the following eight sub-optimization problems.

**Update** $\mathbf{A}_v$: With $\{\mathbf{P}_v, \mathbf{E}_v, \mathbf{W}_v, \mathbf{F}_v, \mathbf{Z}_v\}$ fixed, for each $\mathbf{A}_v$, we need to minimize the following objective function:

$$\left\|\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v - \mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T\right\|_F^2 + \beta \left\|\mathbf{E}_v - \mathbf{A}_v\mathbf{Z}_v\right\|_F^2 \quad \text{s.t.} \quad \mathbf{A}_v^T\mathbf{A}_v = \mathbf{I}. \tag{7}$$

The optimal solution of Eq. (7) is $\mathbf{U}[\mathbf{I},\mathbf{0}]\mathbf{V}^T$, where $\mathbf{U}$ and $\mathbf{V}$ are the left and right singular vectors obtained from the Singular Value Decomposition (SVD) of $(\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v)\mathbf{F}_v\mathbf{W}_v^T + \beta\mathbf{E}_v\mathbf{Z}_v^T$.

**Update** $\mathbf{W}_v$: With $\{\mathbf{P}_v, \mathbf{E}_v, \mathbf{A}_v, \mathbf{F}_v, \mathbf{Z}_v\}$ fixed, for each $\mathbf{W}_v$, we need to minimize the following objective function:

$$\left\|\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v - \mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T\right\|_F^2 \quad \text{s.t.} \quad \mathbf{W}_v^T\mathbf{1} = \mathbf{1}, \mathbf{W}_v \geq 0. \tag{8}$$

By removing the normalization constraint, the partial derivation of Eq. (8) with respect to $\mathbf{W}_v$ is

$$\bar{\mathbf{W}}_v = \mathbf{A}_v^T(\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v)\mathbf{F}_v(\mathbf{F}_v^T\mathbf{F}_v)^{-1}. \tag{9}$$

Then $\mathbf{W}_v$ can be obtained by

$$\min_{\mathbf{W}_v\geq 0, \mathbf{W}_v=1} \|\mathbf{W}_v - \bar{\mathbf{W}}_v\|_F^2 \tag{10}$$

Eq. (10) is an euclidean projection problem on the simplex space [Wang *et al.*, 2020a].

**Update $\mathbf{F}_v$**: With $\{\mathbf{P}_v, \mathbf{E}_v, \mathbf{A}_v, \mathbf{W}_v, \mathbf{Z}_v, \boldsymbol{\mathcal{J}}\}$ fixed, for each $\mathbf{F}_v$, we need to minimize the following objective function:

$$\begin{aligned} & a_v^2 \left\|\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v - \mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T\right\|_F^2 \\ & + \frac{\rho}{2}\left\|\mathbf{F}_v - \mathbf{J}_v + \frac{\mathbf{Y}_v}{\rho}\right\|_F^2 \\ & \text{s.t.} \quad \mathbf{F}_v \geq 0, \end{aligned} \tag{11}$$

The optimal solution of Eq. (11) is

$$((a_v^2(\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v)^T\mathbf{A}_v\mathbf{W}_v + \frac{\rho}{2}(\mathbf{J}_v - \frac{\mathbf{Y}_v}{\rho}))\mathbf{B}_v^{-1}, 0)_+ \tag{12}$$

where $\mathbf{B}_v = a_v^2\mathbf{W}_v^T\mathbf{W}_v + \frac{\rho}{2}\mathbf{I}$.

**Update $\boldsymbol{\mathcal{J}}$**: With $\{\mathbf{P}_v, \mathbf{E}_v, \mathbf{A}_v, \mathbf{W}_v, \mathbf{Z}_v, \mathbf{F}_v\}$ fixed, for $\boldsymbol{\mathcal{J}}$, we need to minimize the following objective function:

$$\psi\|\boldsymbol{\mathcal{J}}\|_{\circledS}^p + \frac{\rho}{2}\left\|\boldsymbol{\mathcal{J}} - (\boldsymbol{\mathcal{F}} + \frac{\boldsymbol{\mathcal{Y}}}{\rho})\right\|_F^2, \tag{13}$$

which has a closed-form solution as Lemma 1 [Gao *et al.*, 2020].

*LEMMA 1. Let $\boldsymbol{\mathcal{T}} \in \mathbb{R}^{n_1\times n_2\times n_3}$ has a tensor-Singular Value Decomposition $\boldsymbol{\mathcal{T}} = \boldsymbol{\mathcal{U}} * \boldsymbol{\Sigma} * \boldsymbol{\mathcal{V}}^\top$, then the optimal solution for*

$$\min_J \frac{1}{2}\|\boldsymbol{\mathcal{H}} - \boldsymbol{\mathcal{T}}\|_F^2 + \tau\|\boldsymbol{\mathcal{H}}\|_{\circledS}^p, \tag{14}$$

*is $\boldsymbol{\mathcal{H}} = \Gamma_\tau(\boldsymbol{\mathcal{T}}) = \boldsymbol{\mathcal{U}} * \text{ifft}\left(P_\tau(\overline{\boldsymbol{\Sigma}})\right) * \boldsymbol{\mathcal{V}}^\top$, where $P_\tau(\overline{\boldsymbol{\Sigma}})$ is an f-diagonal third-order tensor in Fourier domain, whose diagonal elements can be found by using the GST algorithm introduced in* [Gao *et al.*, 2020]. *ifft$(\cdot)$ is the inverse discrete Fourier transform along the third dimension.*

**Update $\mathbf{E}_v$**: With $\{\mathbf{P}_v, \mathbf{A}_v, \mathbf{W}_v, \mathbf{Z}_v, \mathbf{F}_v\}$ fixed, for each $\mathbf{E}_v$, we need to minimize the following objective function:

$$\left\|\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v - \mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T\right\|_F^2 + \beta\|\mathbf{E}_v - \mathbf{A}_v\mathbf{Z}_v\|_F^2 \tag{15}$$

The partial derivation of Eq. (15) with respect to $\mathbf{E}_v$ is

$$\mathbf{E}_v = \frac{(\mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T - \mathbf{P}_v\mathbf{X}_v)\mathbf{G}_v^T + \beta\mathbf{A}_v\mathbf{Z}_v}{1+\beta}, \tag{16}$$

**Update $\mathbf{Z}_v$**: With $\{\mathbf{E}_v, \mathbf{A}_v\}$ fixed, for each $\mathbf{Z}_v$, we need to minimize the following objective function:

$$\frac{1}{2}\|\mathbf{Z}_v - \mathbf{A}_v\mathbf{E}_v\|_F^2 + \frac{1}{2\beta}\|\mathbf{Z}_v\|_1, \tag{17}$$

The optimal solution can be obtained using the well-known soft-threshold operator [Yang and Zhang, 2011].

**Algorithm 1** JVILG

**Input:** $V$-view incomplete dataset $\{\mathbf{X}^v\}_{v=1}^m$, the number of clusters $k$, and the number of anchors $b$.
**Output:** Clustering labels of data points.

```
Initialize: E_v = A_v = W_v = P_v = 0, F_v = [I, 0],
Initialize Y = 0, a_v = 1/m, ρ = 10^-5, μ = 2
while not converge do
    Update A^v by Eq. (7);
    Update W^v by Eq. (10);
    Update F^v by Eq. (12);
    Update J by solving Eq. (13);
    Update E^v by solving Eq. (16);
    Update Z^v by solving Eq. (17);
    Update P^v by solving Eq. (18);
    Update a_v by Eq. (20);
    Update Y and ρ: Y = Y + ρ(F − J), ρ = min(μρ, 10^12);
end while
Calculate the clustering results by using
F = Σ_{v=1}^V a_v^2 F_v / Σ_{v=1}^V a_v^2.
return Clustering results (The label for the corresponding sample is given by the index of the largest element in each column of F).
```

**Update $\mathbf{P}_v$**: With $\{\mathbf{E}_v, \mathbf{A}_v, \mathbf{W}_v, \mathbf{F}_v\}$ fixed, for each $\mathbf{P}_v$, we need to minimize the following objective function:

$$\left\|\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v - \mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T\right\|_F^2 \text{ s.t. } \mathbf{P}_v^T\mathbf{P}_v = \mathbf{I}. \tag{18}$$

The optimal solution of Eq. (18) is $\mathbf{U}_p[\mathbf{I}, \mathbf{0}]\mathbf{V}_p^T$, where $\mathbf{U}_p$ and $\mathbf{V}_p$ are the left and right singular vectors obtained from the Singular Value Decomposition of $(\mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T - \mathbf{E}_v\mathbf{G}_v)\mathbf{X}_v^T$.

**Update $a_v$**: With $\{\mathbf{P}_v, \mathbf{E}_v, \mathbf{A}_v, \mathbf{W}_v, \mathbf{F}_v\}$ fixed, for each $a_v$, we need to minimize the following objective function:

$$\sum_{v=1}^m a_v^2 r_v^2, \quad \text{s.t.} \quad \mathbf{a1} = 1, a_v \geq 0. \tag{19}$$

where $r_v^2 = \left\|\mathbf{P}_v\mathbf{X}_v + \mathbf{E}_v\mathbf{G}_v - \mathbf{A}_v\mathbf{W}_v\mathbf{F}_v^T\right\|_F^2 + \|\mathbf{Z}_v\|_1 + \beta\|\mathbf{E}_v - \mathbf{A}_v\mathbf{Z}_v\|_F^2$. Based on Cauchy-Schwarz inequality, Eq. (19) has a close-form solution as follows:

$$a_v = \frac{\frac{1}{r_v^2}}{\sum_{v=1}^m \frac{1}{r_v^2}} \tag{20}$$

In summary, we outline the entire optimization procedure for solving Eq. (6) in Algorithm 1.

### 3.5 Convergence and Complexity

**Convergence Analysis**: Our problem is bounded due to the summation of norms with positive penalty parameters. Furthermore, the optimization procedure in Algorithm 1 consists of eight sub-problems, each of which can reach exact minimum points. This implies that each sub-problem exhibits a monotonic decrease. Therefore, our algorithm can at least find a locally optimal solution according to the convergence theorem in [Rudin and others, 1964].

**Complexity Analysis**: Our algorithm mainly focuses on solving $\mathbf{A}_v$, $\mathbf{W}_v$, $\mathbf{F}_v$, $\mathcal{J}$, $\mathbf{E}_v$, $\mathbf{Z}_v$, $\mathbf{P}_v$, and $a_v$. The computational complexity in updating these variables iteratively are $\mathcal{O}(mlb^2)$, $\mathcal{O}(mkb)$, $\mathcal{O}(nl\sum_{v=1}^{m} d_v)$, $\mathcal{O}(2mnklog(mk)+m^2kn)$, $\mathcal{O}(nl\sum_{v=1}^{m} d_v)$, $\mathcal{O}(bl\sum_{v=1}^{m} n_v)$, $\mathcal{O}(nl\sum_{v=1}^{m} d_v)$, and $\mathcal{O}(1)$. Due to $b \ll n$, $d_v \ll n$, and $m$, $k$, $l$ are small constants, the main computational complexity of solving Eq. (6) is $\mathcal{O}(max(nl\sum_{v=1}^{m} d_v, m^2kn))$. This indicates that our approach can scale well with data size.

## 4 Experiments

**Datasets:** We evaluate the performance of our method on six extensively used multi-view datasets. ProteinFold [Wan *et al.*, 2024b], Caltech101-7 [Fei-Fei *et al.*, 2006], BDGP [Wang *et al.*, 2022], CCV [Jiang *et al.*, 2011], Animal [Lampert *et al.*, 2009], and NUSWIDEOBJ [Chua *et al.*, 2009]. The sample sizes range from 694 to 30000, and the number of views ranges from 3 to 12. The details of these datasets are shown in Table 1. We follow previous works [Wang *et al.*, 2022; Liu *et al.*, 2022] to construct incomplete multi-view datasets. We set a missing rate $r \in \{10\% : 10\% : 90\%\}$ and randomly select $r$ samples as incomplete data by dropping partial views. The remaining $1 - r$ samples are kept complete.

| Dataset | Size | Classes | Views |
|---|---|---|---|
| ProteinFold | 694 | 27 | 27 × 12views |
| Caltech101-7 | 1474 | 7 | 48/40/254/1984/512/928 |
| BDGP | 2500 | 5 | 1000/500/250 |
| CCV | 6773 | 20 | 20/20/20 |
| Animal | 11673 | 20 | 2689/2000/2001/2000 |
| NUSWIDEOBJ | 30000 | 31 | 65/226/145/74/129 |

Table 1: Incomplete multi-view datasets in experiments.

**Compared Methods:** We compare our approach against two baseline methods and nine state-of-the-art methods: BSV (k-means on the best view), Concate (k-means on concated views), DAIMC [Hu and Chen, 2018], sFSR-IMVC [Long *et al.*, 2023], FIMVC-VIA [Liu *et al.*, 2022], IMVC-CBG [Wang *et al.*, 2022], SMVC-SA [Wen *et al.*, 2023b], PSIMVC-PG [Li *et al.*, 2024], FCMVC-IV [Wan *et al.*, 2024b], DIVIDE [Lu *et al.*, 2024], CPSPAN [Jin *et al.*, 2023]. For the baseline methods BSV and Concate, we use the average feature values of the existing instances in current view to impute the missing instances. For the other compared methods, the hyper-parameters are set to the recommended values in their original papers. Shallow models are performed on a computer with a 3.5GHz AMD Ryzen 9 3950X CPU and 64GB RAM, MATLAB2022b (64-bit).

To evaluate the performance, three commonly used clustering metrics: accuracy (ACC), normalized mutual information (NMI), and Purity are used. The higher the values of these metrics, the better the clustering performance.

### 4.1 Performance Comparison

The clustering results (Each experiment is repeated 10 times.) on four datasets with 10%, 50%, and 90% missing ratios are reported in Table 2. We can draw the following observations:

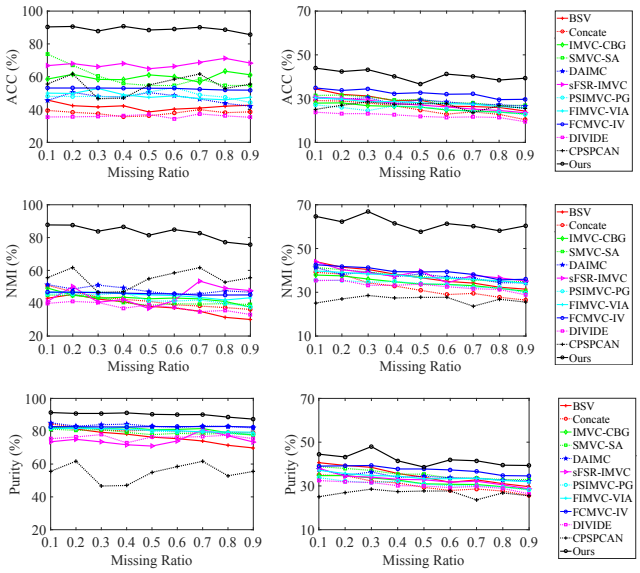



Figure 1: The clustering results for the ACC, NMI, and Purity metrics on two datasets with varying missing ratios. The left column is the Caltech101-7 dataset and the right column is the ProteinFold dataset.

1. Our JVILG achieves the best performance across the three metrics in most cases, demonstrating its superiority in incomplete multi-view clustering tasks. *e.g.*, on the CCV dataset with a 10% missing ratio, our approach outperforms the second-best method, DIVIDE (a deep model) by about 13.73% on the ACC metric.

2. In most cases, an intuitive observation is that the model performance gradually decreases as the view missing ratios increase. The missing view imputation capability of our JVILG model can alleviate the effects of information loss, and help the model maintain the excellent performance. For example, On the BDGP, CCV, and Animal datasets, the ACC values increase by 11.48%, 7.68%, and 4.81%, respectively, as the missing ratio increases from 10% to 50%. Moreover, as the missing ratio reaches 90%, the model performance is still encouraging.

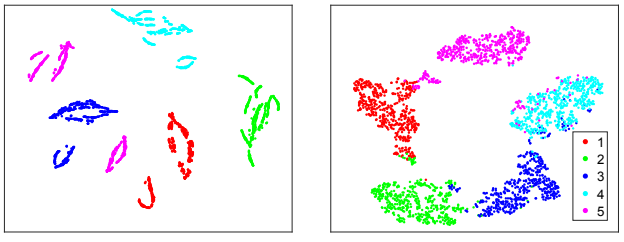
3. Compared to deep models like DIVIDE and CPSPAN, JVILG demonstrates better performance, highlighting the potential of anchors in capturing data distribution.

To further verify the effectiveness of our model, we vary the missing ratios from 10% to 90% with an interval 10% on the ProteinFold and Caltech101-7 datasets. The result curves are shown in Figure 1. We can observe that our method significantly outperforms other compared methods, and its performance remains stable even as the missing ratio increases.

We employed the t-SNE technique [Van der Maaten and Hinton, 2008] to plot the learned cluster representations of our model and the suboptimal sFSR-IMVC model on the BDGP dataset, as shown in Figure 2. Our model demonstrates more clearer cluster discriminability.

| p | Datasets | BDGP | | | CCV | | | Animal | | | NUSWIDEOBJ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Method\Metrics | ACC | NMI | Purity | ACC | NMI | Purity | ACC | NMI | Purity | ACC | NMI | Purity |
| 10% | BSV | 42.56 | 27.08 | 44.67 | 18.24 | 16.65 | 21.75 | 15.30 | 11.99 | 14.12 | 12.22 | 10.86 | 23.81 |
| | Concate | 43.92 | 21.33 | 43.92 | 16.75 | 13.48 | 20.80 | 16.84 | 14.74 | 15.54 | 15.15 | 15.07 | **26.50** |
| | IMVC-CBG | 44.74 | 25.18 | 45.25 | 19.29 | 16.20 | 22.82 | 16.74 | 12.71 | 17.89 | 15.38 | 12.79 | 23.75 |
| | SMVC-SA | 52.63 | 28.68 | 52.95 | 21.65 | 18.17 | 25.00 | 18.11 | 13.51 | 17.66 | 15.58 | 13.71 | 25.05 |
| | DAIMC | 45.55 | 19.21 | 46.17 | 17.04 | 13.21 | 19.55 | 15.04 | 11.94 | 16.71 | 14.61 | 12.91 | 24.26 |
| | sFSR-IMVC | 67.41 | 64.89 | 67.4 | 15.99 | 11.83 | 19.84 | 49.38 | 53.15 | 34.62 | 14.06 | 2.44 | 14.93 |
| | PSIMVC-PG | 49.21 | 25.55 | 50.23 | 17.93 | 13.35 | 20.90 | 16.01 | 12.33 | 17.18 | 12.07 | 10.58 | 22.07 |
| | FIMVC-VIA | 47.83 | 23.48 | 47.83 | 21.91 | 17.43 | 25.36 | 17.40 | 12.52 | 17.98 | 13.50 | 11.72 | 22.75 |
| | FCMVC-IV | 31.96 | 10.53 | 33.60 | 16.65 | 11.70 | 20.49 | 15.55 | 11.03 | 15.70 | 11.57 | 9.30 | 22.02 |
| | DIVIDE | 35.01 | 13.58 | 36.82 | 22.04 | 18.57 | 25.09 | 16.10 | 12.14 | 17.00 | 12.26 | 11.58 | 21.80 |
| | CPSPAN | 46.12 | 22.08 | 46.68 | 17.91 | 15.02 | 21.69 | 13.04 | 8.46 | 15.33 | 10.15 | 9.65 | 21.11 |
| | JVILG | **75.64** | **77.43** | **76.88** | **35.77** | **49.30** | **35.77** | **53.45** | **73.36** | **56.63** | **20.28** | **15.53** | 24.65 |
| 50% | BSV | 37.62 | 22.00 | 38.34 | 16.59 | 13.97 | 19.42 | 14.44 | 9.72 | 16.76 | 11.55 | 9.14 | 21.05 |
| | Concate | 35.82 | 12.99 | 36.33 | 13.60 | 10.63 | 18.09 | 15.61 | 11.51 | 18.17 | 12.63 | 11.93 | 23.57 |
| | IMVC-CBG | 34.44 | 11.14 | 34.52 | 16.34 | 13.30 | 19.61 | 15.05 | 11.04 | 18.36 | 14.69 | 9.82 | 21.72 |
| | SMVC-SA | 51.44 | 22.76 | 51.44 | 19.67 | 15.04 | 22.53 | 15.91 | 11.76 | 18.99 | 15.37 | 11.33 | 23.40 |
| | DAIMC | 28.59 | 5.95 | 29.21 | 13.93 | 8.36 | 17.21 | 15.05 | 11.18 | 18.49 | 13.84 | 10.14 | 21.76 |
| | sFSR-IMVC | 65.86 | 65.71 | 65.92 | 15.43 | 8.77 | 17.54 | 44.42 | 45.83 | 46.97 | 14.91 | 3.11 | 15.48 |
| | PSIMVC-PG | 42.24 | 19.82 | 44.19 | 15.73 | 10.96 | 19.37 | 14.45 | 10.19 | 18.42 | 10.83 | 9.63 | 21.52 |
| | FIMVC-VIA | 38.39 | 13.57 | 39.33 | 19.18 | 14.82 | 22.72 | 16.21 | 11.22 | 19.42 | 12.73 | 10.51 | 22.04 |
| | FCMVC-IV | 34.00 | 10.42 | 35.16 | 16.33 | 11.17 | 19.92 | 14.14 | 8.46 | 17.01 | 10.03 | 6.69 | 19.53 |
| | DIVIDE | 34.40 | 11.97 | 36.28 | 19.31 | 16.01 | 22.86 | 16.33 | 11.87 | 19.69 | 11.87 | 11.29 | 21.91 |
| | CPSPAN | 43.64 | 18.26 | 45.00 | 18.04 | 14.95 | 21.33 | 13.70 | 9.19 | 15.95 | 10.41 | 8.97 | 20.67 |
| | JVILG | **87.12** | **85.84** | **87.12** | **43.45** | **51.87** | **43.45** | **58.26** | **70.97** | **60.08** | **20.01** | **14.00** | **23.80** |
| 90% | BSV | 32.86 | 19.04 | 34.41 | 15.30 | 11.46 | 17.48 | 12.86 | 7.12 | 14.12 | 11.74 | 7.22 | 17.56 |
| | Concate | 32.14 | 10.40 | 32.61 | 12.05 | 9.18 | 16.54 | 13.11 | 9.54 | 15.54 | 10.59 | 9.08 | 21.02 |
| | IMVC-CBG | 32.22 | 10.85 | 33.82 | 14.69 | 10.56 | 18.39 | 14.51 | 9.87 | 17.89 | 14.54 | 0.87 | 20.52 |
| | SMVC-SA | 43.05 | 16.76 | 44.14 | 17.76 | 12.56 | 20.96 | 14.56 | 9.96 | 17.66 | 14.80 | 8.93 | 21.10 |
| | DAIMC | 32.13 | 8.48 | 33.08 | 12.85 | 6.42 | 16.05 | 13.85 | 9.21 | 16.71 | 11.84 | 7.14 | 18.95 |
| | sFSR-IMVC | 63.13 | 59.92 | 63.38 | 13.80 | 9.17 | 17.37 | 34.14 | 32.24 | 34.62 | 14.41 | 5.51 | 16.84 |
| | PSIMVC-PG | 35.64 | 13.68 | 38.80 | 14.25 | 8.93 | 17.51 | 13.99 | 9.19 | 17.18 | 11.10 | 8.56 | 20.56 |
| | FIMVC-VIA | 38.77 | 15.62 | 39.28 | 17.85 | 12.45 | 20.71 | 15.12 | 10.24 | 17.98 | 12.40 | 9.45 | **21.61** |
| | FCMVC-IV | 38.00 | 12.94 | 39.12 | 16.76 | 11.59 | 20.51 | 12.79 | 6.85 | 15.70 | 8.64 | 5.01 | 18.00 |
| | DIVIDE | 31.79 | 6.73 | 33.84 | 16.27 | 11.50 | 19.66 | 13.81 | 9.55 | 17.00 | 11.57 | 10.29 | 20.90 |
| | CPSPAN | 29.12 | 3.68 | 30.36 | 16.86 | 13.34 | 20.58 | 12.59 | 8.58 | 15.33 | 10.60 | 9.06 | 20.85 |
| | JVILG | **73.48** | **75.47** | **74.44** | **35.86** | **44.70** | **37.12** | **49.15** | **67.74** | **56.63** | **18.22** | **12.32** | 21.10 |

Table 2: Results on four datasets with 10%, 50%, and 90% missing ratios. The 1st/2nd best results are marked in **bold** and underline.

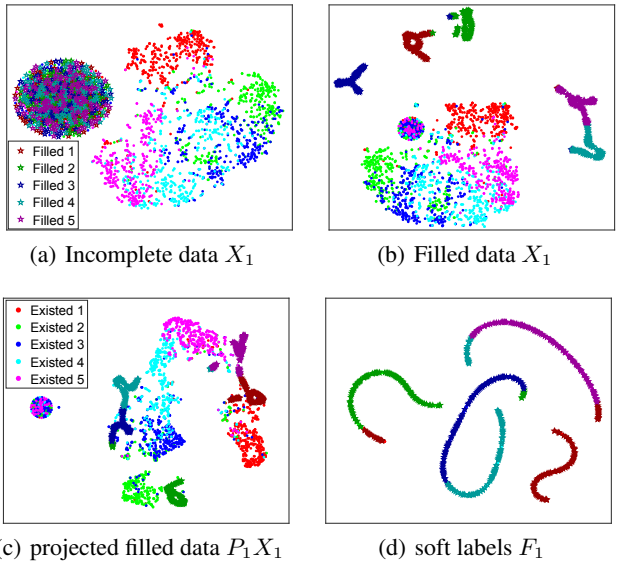



(a) sFSR-IMVC (b) Ours

Figure 2: t-SNE visualization of the learned cluster representations obtained by our JVILG model and sFSR-IMVC method on the BDGP dataset with 10 % missing ratio. Note that 9 clusters appear in (a), while our method identifies exactly 5 clusters in (b).

## 4.2 View Imputation Validation

To validate the missing view imputation capability of our JVILG model, we presented the incomplete data, filled data, projected filled data, and learned soft labels on View #1 of the BDGP dataset with a 50% missing ratio, as shown in Figure 3. Solid circles represent the observed data, while pentagrams denote the filled data. Comparing Figure 3 (a) and (b), the zero imputation results are clustered together, while the imputation values learned by our model are distinctly separated into their respective clusters. Figure 3 (b) and (c) show that the imputation values are effectively integrated into the corresponding clusters through projection learning. This indicates the improved cluster structures in the learned latent representations. Finally, Figure 3 (d) illustrates that the soft labels for the filled data accurately recover the underlying data distribution. These recovered data reside on the same subspace




(a) Incomplete data $X_1$ (b) Filled data $X_1$

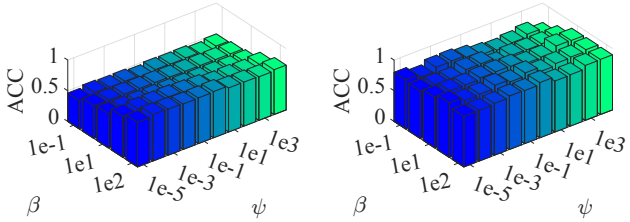
(c) projected filled data $P_1X_1$ (d) soft labels $F_1$

Figure 3: t-SNE visualization of incomplete View #1 data, filled View #1 data, projected filled View #1 data, and learned soft labels on View #1 of the BDGP dataset ($r = 50\%$). Solid circles represent the observed data, while pentagrams denote the filled missing data. Note that the learned soft labels of the imputed views effectively align to the observed views on subspace manifold structures in (d).

manifold structures as the observed data, demonstrating the effectiveness of our JVILG model.

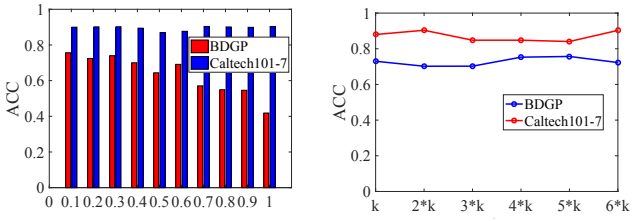
### 4.3 Parameter Study

Our method involves four parameters to be set appropriately, *i.e.*, parameter $\beta$, $\psi$, tensor low rank coefficient $p$, and the number of anchors $b$. As shown in Figure 4, we tune $\beta$ in the range of $10^{[-5:1:4]}$, and vary $\psi$ within $\{0.1, 1, 10, 50, 100\}$. Our JVILG achieves satisfactory performance in a wide scope of $\beta$ and $\psi$. Fine tuning is still necessary for different datasets due to their diverse properties. Figure 5 demonstrates the effects of parameters $p$ and $b$, which are varied as $p \in \{0.1 : 0.1 : 1\}$ and $b \in \{k : k : 6k\}$. Notably, parameter $p$ does impact the model performance, and a small number of anchors can achieve effective and stable results.



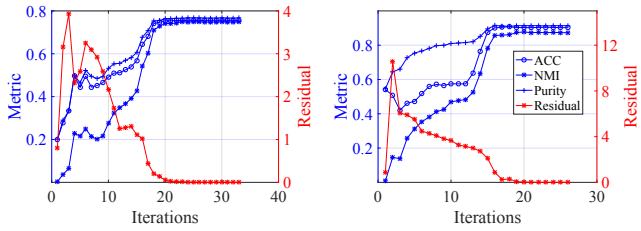

(a) BDGP (b) Caltech101-7

Figure 4: ACC metrics with respect to $\beta$ and $\psi$ on the BDGP and Caltech101-7 datasets ($r = 10\%$).




(a) p (b) b

Figure 5: ACC metrics with respect to tensor low rank coefficient $p$ and anchor number $b$ on the BDGP and Caltech101-7 datasets.




(a) BDGP (b) Caltech101-7

Figure 6: Convergence curves of our method on the BDGP and Caltech101-7 datasets ($r = 10\%$).

### 4.4 Convergence Study

We optimize the objective function iteratively by introducing an auxiliary variable $\mathcal{J}$. Convergence of our algorithm is determined by checking the difference between $\mathcal{F}$ and $\mathcal{J}$. As shown in Figure 6, the difference decreases significantly as the iterations approach 20, ultimately nearing zero. Additionally, clustering metrics such as ACC, NMI, and Purity improve gradually and stabilize as the algorithm converges. This indicates the strong clustering performance of our approach.

## 5 Conclusion

This paper concerned missing view imputation and direct label generation for incomplete multi-view data. Our JVILG model learned versatile anchors to achieve these tasks simultaneously. These anchors can relax the correlations between missing views and semantic clusters, allowing for the imputation of missing view representations and label generation in different feature spaces. For label generation, we introduced membership matrices to learn semantic cluster centroids from fine-grained anchors, from which soft labels are generated to effectively partition the data. Moreover, the tensor Schatten p-norm was imposed on the constructed label tensor to capture complementary information among views. For missing view imputation, we applied subspace correlation to recover missing data from anchors, leveraging their diverse representations. An alternating optimization algorithm was proposed to solve the unified objective function. Extensive experiments demonstrated the effectiveness of our method.

## Acknowledgments

This work was funded by the Science and Technology Development Fund, Macau SAR (File no. 0049/2022/A1, 0050/2024/AGJ), by the University of Macau Development Foundation (File no. MYRG-GRG2024-00181-FST-UMDF).

## References

[Bai *et al.*, 2024] Gairui Bai, Wei Xi, and Jialin Zhuo. Adapting multi-view correlation learning for enhanced cervical dysplasia diagnosis. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 1796–1801. IEEE, 2024.

[Boyd *et al.*, 2011] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.

[Cai *et al.*, 2013] Xiao Cai, Feiping Nie, and Heng Huang. Multi-view k-means clustering on big data. In *Twenty-Third International Joint conference on artificial intelligence*, 2013.

[Chen *et al.*, 2023] 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*, 2023.

[Chen *et al.*, 2024] Man-Sheng Chen, Chang-Dong Wang, Dong Huang, Jian-Huang Lai, and S Yu Philip. Concept factorization based multiview clustering for large-scale data. *IEEE Transactions on Knowledge and Data Engineering*, 2024.

[Chua *et al.*, 2009] Tat-Seng Chua, Jinhui Tang, Richang Hong, Haojie Li, Zhiping Luo, and Yantao Zheng. Nus-wide: a real-world web image database from national university of singapore. In *Proceedings of the ACM international conference on image and video retrieval*, pages 1–9, 2009.

[Cui *et al.*, 2023] Chenhang Cui, Yazhou Ren, Jingyu Pu, Xiaorong Pu, and Lifang He. Deep multi-view subspace clustering with anchor graph. *arXiv preprint arXiv:2305.06939*, 2023.

[Fei-Fei *et al.*, 2006] Li Fei-Fei, Robert Fergus, and Pietro Perona. One-shot learning of object categories. *IEEE transactions on pattern analysis and machine intelligence*, 28(4):594–611, 2006.

[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.

[Guo and Ye, 2019] Jun Guo and Jiahui Ye. Anchors bring ease: An embarrassingly simple approach to partial multi-view clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 118–125, 2019.

[Hu and Chen, 2018] Menglei Hu and Songcan Chen. Doubly aligned incomplete multi-view clustering. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 2262–2268, 2018.

[Hu and Chen, 2019] Menglei Hu and Songcan Chen. One-pass incomplete multi-view clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 3838–3845, 2019.

[Jiang *et al.*, 2011] Yu-Gang Jiang, Guangnan Ye, Shih-Fu Chang, Daniel Ellis, and Alexander C Loui. Consumer video understanding: A benchmark database and an evaluation of human and machine performance. In *Proceedings of the 1st ACM International Conference on Multimedia Retrieval*, pages 1–8, 2011.

[Jin *et al.*, 2023] Jiaqi Jin, Siwei Wang, Zhibin Dong, Xinwang Liu, and En Zhu. Deep incomplete multi-view clustering with cross-view partial sample and prototype alignment. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11600–11609, 2023.

[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. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 4412–4419, 2020.

[Lampert *et al.*, 2009] Christoph H Lampert, Hannes Nickisch, and Stefan Harmeling. Learning to detect unseen object classes by between-class attribute transfer. In *2009 IEEE conference on computer vision and pattern recognition*, pages 951–958. IEEE, 2009.

[Li *et al.*, 2023a] Xingfeng Li, Yinghui Sun, Quansen Sun, Jia Dai, and Zhenwen Ren. Distribution consistency based fast anchor imputation for incomplete multi-view clustering. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 368–376, 2023.

[Li *et al.*, 2023b] Xingfeng Li, Yinghui Sun, Quansen Sun, Zhenwen Ren, and Yuan Sun. Cross-view graph matching guided anchor alignment for incomplete multi-view clustering. *Information Fusion*, 100:101941, 2023.

[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.

[Liu *et al.*, 2021a] Jiyuan Liu, Xinwang Liu, Yi Zhang, Pei Zhang, Wenxuan Tu, Siwei Wang, Sihang Zhou, Weixuan Liang, Siqi Wang, and Yuexiang Yang. Self-representation subspace clustering for incomplete multi-view data. In *Proceedings of the 29th ACM international conference on multimedia*, pages 2726–2734, 2021.

[Liu *et al.*, 2021b] Xinwang Liu, Li Liu, Qing Liao, Siwei Wang, Yi Zhang, Wenxuan Tu, Chang Tang, Jiyuan Liu, and En Zhu. One pass late fusion multi-view clustering. In *International conference on machine learning*, pages 6850–6859. PMLR, 2021.

[Liu *et al.*, 2022] Suyuan Liu, Xinwang Liu, Siwei Wang, Xin Niu, and En Zhu. Fast incomplete multi-view clustering with view-independent anchors. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[Long *et al.*, 2023] 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*, 2023.

[Lu *et al.*, 2024] Yiding Lu, Yijie Lin, Mouxing Yang, Dezhong Peng, Peng Hu, and Xi Peng. Decoupled contrastive multi-view clustering with high-order random walks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 14193–14201, 2024.

[Rudin and others, 1964] Walter Rudin et al. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1964.

[Tang *et al.*, 2022] Chang Tang, Zhenglai Li, Jun Wang, Xinwang Liu, Wei Zhang, and En Zhu. Unified one-step multi-view spectral clustering. *IEEE Transactions on Knowledge and Data Engineering*, 35(6):6449–6460, 2022.

[Van der Maaten and Hinton, 2008] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[Wan *et al.*, 2023] Xinhang Wan, Xinwang Liu, Jiyuan Liu, Siwei Wang, Yi Wen, Weixuan Liang, En Zhu, Zhe Liu, and Lu Zhou. Auto-weighted multi-view clustering for large-scale data. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 10078–10086, 2023.

[Wan *et al.*, 2024a] Xinhang Wan, Jiyuan Liu, Xinbiao Gan, Xinwang Liu, Siwei Wang, Yi Wen, Tianjiao Wan, and En Zhu. One-step multi-view clustering with diverse representation. *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

[Wan *et al.*, 2024b] Xinhang Wan, Bin Xiao, Xinwang Liu, Jiyuan Liu, Weixuan Liang, and En Zhu. Fast continual multi-view clustering with incomplete views. *IEEE Transactions on Image Processing*, 2024.

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

[Wang *et al.*, 2020b] Qianqian Wang, Jiafeng Cheng, Quanxue Gao, Guoshuai Zhao, and Licheng Jiao. Deep multi-view subspace clustering with unified and discriminative learning. *IEEE Transactions on Multimedia*, 23:3483–3493, 2020.

[Wang *et al.*, 2021a] Qianqian Wang, Zhengming Ding, Zhiqiang Tao, Quanxue Gao, and Yun Fu. Generative partial multi-view clustering with adaptive fusion and cycle consistency. *IEEE Transactions on Image Processing*, 30:1771–1783, 2021.

[Wang *et al.*, 2021b] 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.

[Wang *et al.*, 2022] Siwei Wang, Xinwang Liu, Li Liu, Wenxuan Tu, Xinzhong Zhu, Jiyuan Liu, Sihang Zhou, and En Zhu. Highly-efficient incomplete large-scale multi-view clustering with consensus bipartite graph. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9776–9785, 2022.

[Wen *et al.*, 2019] 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*, volume 33, pages 5393–5400, 2019.

[Wen *et al.*, 2021a] Jie Wen, Huijie Sun, Lunke Fei, Jinxing Li, Zheng Zhang, and Bob Zhang. Consensus guided incomplete multi-view spectral clustering. *Neural Networks*, 133:207–219, 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.*, 2023a] 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.

[Wen *et al.*, 2023b] Yi Wen, Siwei Wang, Ke Liang, Weixuan Liang, Xinhang Wan, Xinwang Liu, Suyuan Liu, Jiyuan Liu, and En Zhu. Scalable incomplete multi-view clustering with structure alignment. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 3031–3040, 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.

[Yang and Zhang, 2011] Junfeng Yang and Yin Zhang. Alternating direction algorithms for \ell_1-problems in compressive sensing. *SIAM journal on scientific computing*, 33(1):250–278, 2011.

[Yu *et al.*, 2023] Penghang Yu, Zhiyi Tan, Guanming Lu, and Bing-Kun Bao. Multi-view graph convolutional network for multimedia recommendation. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 6576–6585, 2023.

[Zhao *et al.*, 2023] Xiaojia Zhao, Qiangqiang Shen, Yongyong Chen, Yongsheng Liang, Junxin Chen, and Yicong Zhou. Self-completed bipartite graph learning for fast incomplete multi-view clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023.