Sample-Level Cross-View Similarity Learning for Incomplete Multi-View Clustering

Suyuan Liu[1], Junpu Zhang[1], Yi Wen[1], Xihong Yang[1], Siwei Wang[2], Yi Zhang[1], En Zhu[1], Chang Tang[3], Long Zhao[4], Xinwang Liu[1]*

[1] School of Computer, National University of Defense Technology, Changsha, China, 410073
[2] Intelligent Game and Decision Lab, Beijing, China, 100091
[3] School of Computer Science, China University of Geosciences, Wuhan, China, 430074
[4] Shandong Computer Science Center, Qilu University of Technology, Jinan, China, 250000
suyuanliu@nudt.edu.cn, zhangjunpu@nudt.edu.cn, wenyiwy2022@163.com, xihong_edu@163.com, wangsiwei13@nudt.edu.cn, zhangy@nudt.edu.cn, enzhu@nudt.edu.cn, tangchang@cug.edu.cn, Zhaolong@qlu.edu.cn, xinwangliu@nudt.edu.cn

Abstract

Incomplete multi-view clustering has attracted much attention due to its ability to handle partial multi-view data. Recently, similarity-based methods have been developed to explore the complete relationship among incomplete multi-view data. Although widely applied to partial scenarios, most of the existing approaches are still faced with two limitations. Firstly, fusing similarities constructed individually on each view fails to yield a complete unified similarity. Moreover, incomplete similarity generation may lead to anomalous similarity values with column sum constraints, affecting the final clustering results. To solve the above challenging issues, we propose a Sample-level Cross-view Similarity Learning (SCSL) method for Incomplete Multi-view Clustering. Specifically, we project all samples to the same dimension and simultaneously construct a complete similarity matrix across views based on the inter-view sample relationship and the intra-view sample relationship. In addition, a simultaneously learning consensus representation ensures the validity of the projection, which further enhances the quality of the similarity matrix through the graph Laplacian regularization. Experimental results on six benchmark datasets demonstrate the ability of SCSL in processing incomplete multi-view clustering tasks. Our code is publicly available at https://github.com/Tracesource/SCSL.

Introduction

Multi-view clustering (MVC) has garnered considerable attention from researchers as an effective machine learning paradigm for data analysis (Jiang et al. 2022; Wan et al. 2022; Dong et al. 2023). The objective of MVC is to leverage the features of samples across multiple views and partition the data points into a predetermined number of categories, where samples within the same category should be as similar as possible (Wang et al. 2022a; Wen et al. 2023a; Yang et al. 2021). The key challenge in MVC lies in determining the consistent and complementary information across multiple views to achieve clustering performance that surpasses that of single-view clustering. Over the past few decades, numerous MVC methods have been proposed and have demonstrated excellent performance (Zhang et al. 2022; Wen et al. 2023b; Yu et al. 2023).

However, in practical applications, multi-view data often suffers from partial missingness due to limitations in the data acquisition process, noise from collection devices, or other uncontrollable factors (Xu et al. 2019; Wang et al. 2022b). For instance, a patient's medical record may be incomplete because several tests were not conducted, or data from multiple radar sensors may have missing values caused by channel interference. Traditional MVC models are typically built on the fundamental assumption of complete data and are unable to directly handle missingness in multi-view data (Wen et al. 2022). To address these challenges, a series of incomplete multi-view clustering (IMVC) methods have emerged in recent years (Liu et al. 2021; Lin et al. 2021; Liu et al. 2022; He et al. 2023; Yang et al. 2023). Based on the strategies employed to handle missing data, the existing IMVC methods can be categorized into three types: imputation-based, representation-based, and similarity-based.

The imputation-based IMVC methods address incomplete data through heuristic strategies or optimization techniques, followed by fusion and clustering operations on the filling data (Liu et al. 2020a; Xia et al. 2023). Liu et al. further propose integrating missing value imputation and clustering into a unified learning process (Liu et al. 2020b). Based on the assumption that each sample in incomplete multi-view data exists in at least one view, representation-based IMVC methods learn a consistent low-dimensional representation based on the samples present in each view and then apply clustering algorithms to this representation (Lv et al. 2022; Deng et al. 2023). Wen et al. suggest first learning incomplete low-dimensional representations on each view separately and then integrating them into a unified latent representation (Wen et al. 2018). Similarity-based IMVC methods construct partial similarity graphs of each view and then fuse them to build a complete similarity matrix (Liu et al. 2023). Guo et al. propose to average the common similarity parts among views and concatenate the unique parts to obtain a complete similarity matrix (Guo and Ye 2019).

Compared to the previous two methods, the similarity-

*Corresponding author

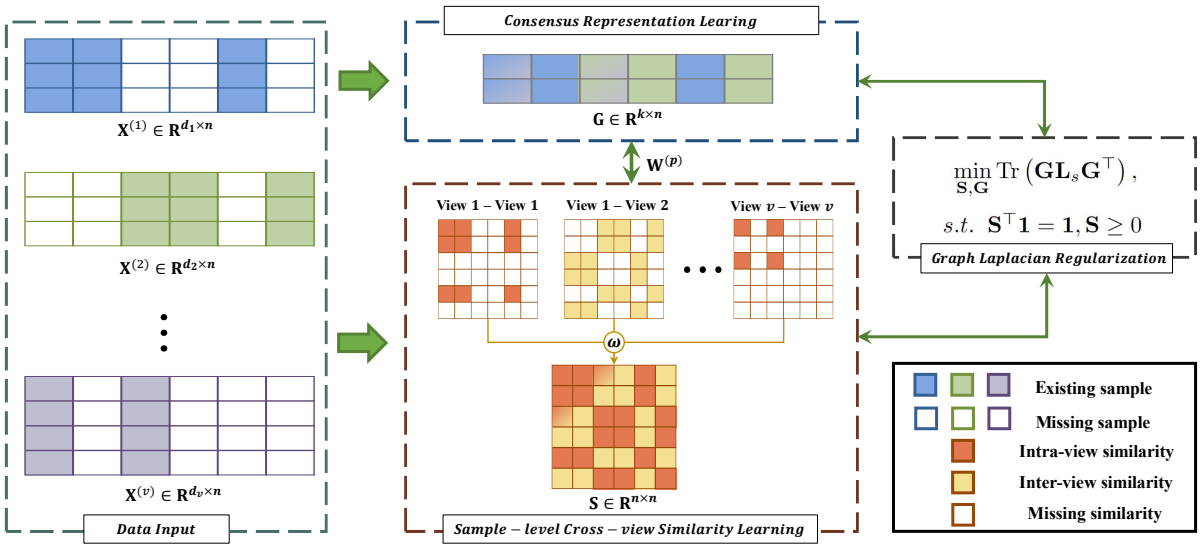


Figure 1: The framework of SCSL.

based approaches avoid the additional computational cost caused by imputation and are not affected by potential noise during dimensionality reduction. Directly targeting the clustering task, similarity construction is a promising method for solving the IMVC problem. However, existing methods have three shortcomings. Firstly, although samples always exist in at least one view, there may be cases where two samples do not appear in the same view simultaneously. In such cases, the similarity matrix constructed in all views will lack the corresponding entry that captures the relationship between these two samples. Existing methods overlook this possibility, resulting in an incomplete consensus similarity matrix. Furthermore, most methods impose a constraint on the similarity matrix, requiring the columns to sum up to 1. When there is a missing entry with a relatively larger similarity, other originally smaller similarities will increase dramatically, leading to abnormal values and affecting the final clustering results. Moreover, the similarity-based methods excessively rely on the quality of the constructed similarity. Merely using similarity as the basis for subsequent clustering processes may result in unsatisfactory outcomes.

To address the issues above, this paper proposes a novel approach named SCSL. Firstly, we suggest constructing the similarity matrix based on the relationships between samples that exist across all views, rather than solely considering intra-view relationships. This avoids the drawbacks of incomplete similarity matrices constructed within individual views. The projection matrices learned on each view ensure that all samples can be measured on the same dimension for distance calculation. To mitigate the impact of imbalanced occurrences of elements on the final similarity construction, we introduce a balancing coefficient to ensure equal contributions from all entries. Additionally, we incorporate a consistent latent representation learning module that enhances the reliability of projections by sharing the same projection as the previous module. The introduction of an additional graph Laplacian term connects the latent representation with the similarity matrix, enabling a global exploration of relationships between samples and further improving the quality of the learned complete similarity matrix. Fig. 1 displays the framework of the presented SCSL.

In summary, the proposed SCSL method in this paper makes several contributions:

- Addressing the issue of incomplete similarity construction, we propose a sample-level cross-view similarity learning approach. To the best of our knowledge, this is the first attempt to construct similarities between samples from a cross-view perspective to tackle the IMVC problem. By simultaneously utilizing intra-view and inter-view sample relationships, this method effectively avoids graph incompleteness and abnormal similarities.
- By learning consistent representations through shared projections across views, the reliability of projections is greatly enhanced. The additional introduction of a graph Laplacian term establishes the connection between consistent representations and the similarity matrix, mutually reinforcing each other and further improving quality.
- We propose a three-step iterative optimization algorithm to solve the optimization problem and prove its convergence. Experimental results on multiple datasets validate the effectiveness of this method.

Related Work

In this section, we provide a brief introduction to IMVC and further focus on the IMVC similarity-based methods. Table 1 presents the symbols used in this paper.

Incomplete Multi-view Clustering

Incomplete multi-view clustering extends MVC algorithms to scenarios with missing data by inferring the missing parts based on the available part or directly constructing a complete unified representation. Existing IMVC approaches can

Notation	Definition
n	Number of samples
n_p	Number of existing samples
k	Number of clusters
v	Number of views
d_p	Dimension
ζ_p	Existing sample indices set
$\mathbf{X}^{(p)} \in \mathbb{R}^{d_p \times n}$	Data matrix
$\mathbf{W}^{(p)} \in \mathbb{R}^{d_p \times k}$	Projection matrix
$\mathbf{H}^{(p)} \in \{0,1\}^{n \times n_p}$	Missing index matrix
$\mathbf{S} \in \mathbb{R}^{n \times n}$	Similarity matrix
$\mathbf{L}_s \in \mathbb{R}^{n \times n}$	Laplacian matrix of $\mathbf{S}$
$\mathbf{G} \in \mathbb{R}^{k \times n}$	Consensus representation matrix

Table 1: Main symbols used throughout the paper.

be divided into three categories based on how they handle missing data: imputation-based methods, representation-based methods, and similarity-based methods.

Imputation-based methods are mainly developed based on multi-kernel clustering, aiming to fill in the missing values with different strategies (Yin and Sun 2022; Xia et al. 2023). The aforementioned methods all follow a two-step process: imputation and then clustering, where the two processes cannot contribute to each other. Liu et al. introduce the incomplete data as variables into the optimization process, integrating imputation and clustering into the same framework (Liu et al. 2017, 2020b).

Representation-based methods assume that views share a low-dimensional subspace, aiming to learn a unified low-dimensional representation for subsequent clustering (Deng et al. 2020; Wen et al. 2020). Shao et al. independently learn latent feature matrices on each view and then fuse them into a consistent matrix (Shao, He, and Philip 2015). Liang et al. propose to learn sample-level weights, further improving clustering performance (Liang, Yang, and Xie 2022).

Similarity-based methods focus on the similarity closely related to the clustering task, aiming to reconstruct a complete similarity graph by leveraging the relationships among samples across different views (Fang et al. 2020; Wen et al. 2021b). Wang et al. are the first to address the incomplete multi-view data problem by transferring it from the data domain to the similarity domain (Wang et al. 2019). In contrast to the above methods, Wen et al. construct local similarity graphs on each view separately and then recover the incomplete parts through a common graph (Wen et al. 2021a).

Similarity-based Incomplete Multi-view Clustering

The similarity-based IMVC methods address the partial multi-view data problem by constructing a similarity matrix among all samples. Specifically, given the multi-view data $\left\{\mathbf{X}^{(p)}\right\}_{p=1}^{v}$, where n is the number of samples and d_p is the dimensionality of samples in each view, the framework of the similarity-based method is as follows:

$$\min_{\mathbf{S}^{(p)},\mathbf{S}} \sum_{p=1}^{v} \sum_{i,j \in \zeta^p} \left\| \mathbf{x}_i^{(p)} - \mathbf{x}_j^{(p)} \right\|_2^2 {s_{ij}^{(p)}}^2 + \alpha f(\mathbf{S}, \mathbf{S}^{(p)}), \quad (1)$$

$$s.t.\ \ {\mathbf{S}^{(p)}}^\top \mathbf{1} = \mathbf{1}, \mathbf{S}^{(p)} \geq 0, \mathbf{S}^\top \mathbf{1} = \mathbf{1}, \mathbf{S} \geq 0,$$

where ζ^p represents the set of sample indices existing in the p-th view. $\mathbf{S}^{(p)}$ denotes the partial similarity matrix constructed on the p-th view, and $\mathbf{S}$ is the consensus matrix obtained through the fusion term f.

Many works have extended the above framework by incorporating different regularization terms into $\mathbf{S}^{(p)}$ or modifying the form of the fusion term f. However, we observe that in missing scenarios, some combinations of sample pairs may not occur in any view. This issue leads to an incomplete $\mathbf{S}$ learned in the end, thereby affecting the subsequent clustering performance. Even worse, the presence of missing elements in the objective function introduces abnormal similarities, where originally larger similarities become zero and smaller similarities exhibit a sharp increase. In the next section, we propose SCSL to address the above problems.

Method

In this section, we begin by introducing each part of the proposed model, followed by a detailed explanation of the optimization procedure. Finally, we delve into the complexity analysis of SCSL. The proposed SCSL consists of three main sub-models: sample-level cross-view similarity learning, consensus representation learning, and graph Laplacian regularization.

Sample-level Cross-view Similarity Learning

Existing similarity-based IMVC methods construct the similarity matrix based on the distances between samples within a single view, leading to a lack of similarity for certain sample pairs. Additionally, the constraints on column sums further affect the construction of the similarity matrix, causing a sudden increase in small similarities in the presence of missing larger values, severely hindering clustering. The key reason for such a problem is that combining the similarity matrices $\mathbf{S}^{(p)} \in \mathbb{R}^{n_p \times n_p}$ constructed for individual samples based on different views does not cover all relationships between samples. For example, in the case of data with two views, if the i-th sample exists in the first view but is missing in the j-th sample, while the opposite is true in the second view, existing methods fail to construct the similarity between the i-th and j-th samples.

To address this issue, we propose utilizing not only the distances between samples within each view but also the distances between samples across different views to construct a comprehensive similarity matrix, denoted as $\mathbf{S} \in \mathbb{R}^{n \times n}$. Since the sample dimensions vary across different views, we introduce a projection matrix $\mathbf{W}^{(p)}$ for each view to project all samples into a common k-dimensional space and compute their distances. As a result, the objective function for sample-level cross-view similarity learning can be formulated as follows:

$$\min_{\mathbf{S},\mathbf{W}^{(p)}} \sum_{p=1}^{v} \sum_{u=1}^{v} \sum_{i \in \zeta^p} \sum_{j \in \zeta^u} \left\| {\mathbf{W}^{(p)}}^\top \mathbf{x}_i^{(p)} - {\mathbf{W}^{(u)}}^\top \mathbf{x}_j^{(u)} \right\|_2^2 s_{ij}^2, \quad (2)$$

$$s.t.\ \ \mathbf{S}^\top \mathbf{1} = \mathbf{1}, \mathbf{S} \geq 0, {\mathbf{W}^{(p)}}^\top \mathbf{W}^{(p)} = \mathbf{I},$$

where ζ_p and ζ_u represent the sets of existing sample indices for the p-th and u-th views, respectively. The orthogonality constraints on the projection matrices reduce redundancy in the projected information. In the original method, the set of constructed sample pairs is $\bigcup_{p=1}^{v} (\zeta_p \times \zeta_p)$, which cannot cover the relationship among all samples. While in our approach, the set is $\bigcup_{p,u=1}^{v} (\zeta_p \times \zeta_u)$, which actually construct a complete similarity. Mathematically,

$$\left| \bigcup_{p=1}^{v} (\zeta_p \times \zeta_p) \right| \leq \left| \bigcup_{p,u=1}^{v} (\zeta_p \times \zeta_u) \right| = n^2. \tag{3}$$

Consensus Representation Learning

To enhance the reliability of projections across different views and explore their consistency, we introduce a module for consistent representation learning. Specifically, we aim to find a common latent representation $\mathbf{G}$ that approximates the existing samples in each view. The objective function can be formulated as follows:

$$\min_{\mathbf{W}^{(p)}, \mathbf{G}} \sum_{p=1}^{v} \left\| \mathbf{X}^{(p)} \mathbf{H}^{(p)} - \mathbf{W}^{(p)} \mathbf{G} \mathbf{H}^{(p)} \right\|_F^2, \tag{4}$$

$$s.t.\ \ {\mathbf{W}^{(p)}}^\top \mathbf{W}^{(p)} = \mathbf{I},$$

where $\mathbf{H}^{(p)}$ is the missing index matrix, let w be an indicator vector containing the indices of n_p existing samples for the p-th view in a specific order, $\mathbf{H}_{ij}^{(p)}$ is set to 1 when w_{pi} equals the value j. It is worth noting that the projection matrix $\mathbf{W}^{(p)}$ in this module remains consistent with the previous one, enabling $\mathbf{G}$ to contribute to the construction of a more reliable projection matrix.

Graph Laplacian Regularization

According to the manifold learning theory, if two data points have high similarity, their reconstructed low-dimensional representations should also be close to each other. Based on this characteristic, we can further guide the learning of similarity using the reconstructed consistent representation. Specifically, we establish a connection between the similarity matrix and the consistent representation through a graph Laplacian regularization term. The objective function can be defined as follows:

$$\min_{\mathbf{S}, \mathbf{G}} \operatorname{Tr}\left(\mathbf{G} \mathbf{L}_s \mathbf{G}^\top \right), \tag{5}$$

$$s.t.\ \ \mathbf{S}^\top \mathbf{1} = \mathbf{1}, \mathbf{S} \geq 0,$$

where $\mathbf{L}_s$ denotes the Laplacian matrix corresponding to $\mathbf{S}$, defined as $\mathbf{L}_s = \mathbf{D} - \frac{\mathbf{S}+\mathbf{S}^\top}{2}$. $\mathbf{D}$ is a diagonal matrix constructed from the rows and columns of $\frac{\mathbf{S}+\mathbf{S}^\top}{2}$. Unlike the local similarity built in the first module, this term directly explores the global structure of the data through consistent representation, thereby enhancing the quality of the similarity matrix.

Overall Model of SCSL

To mitigate the impact of varying occurrence frequencies among sample pairs on the construction of similarity, we assign a normalized weight $\boldsymbol{\omega}$ to each element of the similarity construction. Specifically, $\omega_{ij} = \frac{1}{n_{ij}}$, where n_{ij} denotes the occurrences number of the corresponding sample pair. Finally, by incorporating two hyperparameters to regulate the influence among the three modules, we derive the overall objective function as follows:

$$\begin{aligned} \min_{\mathbf{S}, \mathbf{W}^{(p)}, \mathbf{G}} & \sum_{p=1}^{v} \sum_{u=1}^{v} \sum_{i \in \zeta^p} \sum_{j \in \zeta^u} \omega_{ij} \left\| {\mathbf{W}^{(p)}}^\top \mathbf{x}_i^{(p)} - {\mathbf{W}^{(u)}}^\top \mathbf{x}_j^{(u)} \right\|_2^2 s_{ij}^2 \\ & + \beta \sum_{p=1}^{v} \left\| \mathbf{X}^{(p)} \mathbf{H}^{(p)} - \mathbf{W}^{(p)} \mathbf{G} \mathbf{H}^{(p)} \right\|_F^2 + \lambda \operatorname{Tr}\left(\mathbf{G} \mathbf{L}_s \mathbf{G}^\top \right), \\ & s.t.\ \ \mathbf{S}^\top \mathbf{1} = \mathbf{1}, \mathbf{S} \geq 0, {\mathbf{W}^{(p)}}^\top \mathbf{W}^{(p)} = \mathbf{I}. \end{aligned} \tag{6}$$

Optimization

To tackle the aforementioned optimization problem, a sequential variable updating scheme is employed, whereby each variable is updated while maintaining the others as constants. The specific details are outlined as follows:

(1) **Update S** While keeping the other two variables $\mathbf{W}^{(p)}$ and $\mathbf{G}$ constant, the optimization of $\mathbf{S}$ is attained by solving the subsequent equation.

$$\begin{aligned} \min_{\mathbf{S}} & \sum_{p=1}^{v} \sum_{u=1}^{v} \sum_{i \in \zeta^p} \sum_{j \in \zeta^u} \omega_{ij} \left\| {\mathbf{W}^{(p)}}^\top \mathbf{x}_i^{(p)} - {\mathbf{W}^{(u)}}^\top \mathbf{x}_j^{(u)} \right\|_2^2 s_{ij}^2 \\ & + \lambda \operatorname{Tr}\left(\mathbf{G} \mathbf{L}_s \mathbf{G}^\top \right), \\ & s.t.\ \ \mathbf{S}^\top \mathbf{1} = \mathbf{1}, \mathbf{S} \geq 0. \end{aligned} \tag{7}$$

Utilizing $\mathbf{s}_j$ to denote the j-th column of $\mathbf{S}$, we note the independence of Eq. (7) across distinct values of j, enabling separate solutions for each j:

$$\min_{\mathbf{s}_j} {\mathbf{s}_j}^\top \mathbf{A}^j \mathbf{s}_j + \lambda \mathbf{b}_j^\top \mathbf{s}_j, \tag{8}$$

$$s.t.\ \ {\mathbf{s}_j}^\top \mathbf{1} = 1, \mathbf{s}_j \geq 0,$$

where $\mathbf{A}^j$ is a diagonal matrix with diagonal elements $\mathbf{A}_{ii}^j = \sum_{p=1}^{v} \sum_{u=1}^{v} \omega_{ij} \left\| {\mathbf{W}^{(p)}}^\top \mathbf{x}_i^{(p)} - {\mathbf{W}^{(u)}}^\top \mathbf{x}_j^{(u)} \right\|_2^2$, $b_{ji} = \frac{1}{2} \left\| \mathbf{g}_i - \mathbf{g}_j \right\|_2^2$.

Nie et al. offer a comprehensive delineation of the technique to solve equation Eq. (8) (Nie et al. 2016). The optimal $\mathbf{s}_j$ can be expressed as:

$$s_{ji} = \left(\frac{\eta - \lambda b_{ji}}{2 \mathbf{A}_{ii}^j} \right)_+, \tag{9}$$

where $(\cdot)_+ = max(\cdot, 0)$, and the parameter η is obtained as the root of the equation $g_j(\eta) = 0$, which can be readily computed using Newton's method.

$$g_j(\eta) = \sum_{i=1}^{n} \left(\frac{\eta - \lambda b_{ji}}{2 \mathbf{A}_{ii}^j} \right)_+ - 1. \tag{10}$$

(2) **Update $\mathbf{W}^{(p)}$** While keeping the other two variables $\mathbf{S}$ and $\mathbf{G}$ constant, the optimization of $\mathbf{W}^{(p)}$ is attained by solving the subsequent equation.

$$\min_{\mathbf{W}^{(p)}} \sum_{u=1}^{v} \sum_{i\in\zeta^p} \sum_{j\in\zeta^u} \omega_{ij} \left\| {\mathbf{W}^{(p)}}^\top \mathbf{x}_i^{(p)} - {\mathbf{W}^{(u)}}^\top \mathbf{x}_j^{(u)} \right\|_2^2 (s_{ij}^2 + s_{ji}^2 1_{u\neq p}) + \beta \left\| \mathbf{X}^{(p)}\mathbf{H}^{(p)} - \mathbf{W}^{(p)}\mathbf{G}\mathbf{H}^{(p)} \right\|_F^2, \quad s.t.\ {\mathbf{W}^{(p)}}^\top \mathbf{W}^{(p)} = \mathbf{I}, \tag{11}$$

where $1_{u\neq p}$ denotes an indicator function that evaluates to 1 when variables u and p are not equal, and 0 otherwise. Removing constant terms unrelated to $\mathbf{W}^{(p)}$, Eq. (11) is simplified to the following form,

$$\min_{\mathbf{W}^{(p)}} \operatorname{Tr}\left({\mathbf{W}^{(p)}}^\top \mathbf{C}^{(p)} \mathbf{W}^{(p)} + {\mathbf{W}^{(p)}}^\top \mathbf{D}^{(p)} \right), \quad s.t.\ {\mathbf{W}^{(p)}}^\top \mathbf{W}^{(p)} = \mathbf{I}, \tag{12}$$

where $\mathbf{C}^{(p)} = \mathbf{Q}^{(p)}\mathbf{Z}\left(\mathbf{Q}^{(p)}\right)^\top, \mathbf{Z} = diag\left(\mathbf{Y}\sum_{u=1}^{v}\mathbf{a}^{(u)}\right) - \mathbf{Y}, \mathbf{Y} = \boldsymbol{\omega}\odot\left(\mathbf{S}\odot\mathbf{S} + \mathbf{S}^\top\odot\mathbf{S}^\top\right). \mathbf{a}^{(u)} = \mathbf{H}^{(u)}\mathbf{1}, \mathbf{Q}^{(p)} = \mathbf{X}^{(p)} \odot \mathbf{1}{\mathbf{a}^{(p)}}^\top$. The symbol $\odot$ denotes the Hadamard product. And $\mathbf{D}^{(p)} = -2\sum_{u=1}^{v}\mathbf{Q}^{(p)}\mathbf{Y}\left(\mathbf{Q}^{(u)}\right)^\top\mathbf{W}^{(u)} + \beta\mathbf{Q}^{(p)}\mathbf{G}^\top$. For detailed derivation process regarding $\mathbf{C}^{(p)}$ and $\mathbf{D}^{(p)}$, please refer to the appendix.

Eq.(12) can be efficiently solved using the Generalized Power Iteration Method (GPI) (Nie, Zhang, and Li 2017).

(3) **Update $\mathbf{G}$** While keeping the other two variables $\mathbf{S}$ and $\mathbf{W}^{(p)}$ constant, the optimization of $\mathbf{G}$ is attained by solving the subsequent equation.

$$\min_{\mathbf{G}} \sum_{p=1}^{v} \beta \left\| \mathbf{X}^{(p)}\mathbf{H}^{(p)} - \mathbf{W}^{(p)}\mathbf{G}\mathbf{H}^{(p)} \right\|_F^2 + \lambda \operatorname{Tr}\left(\mathbf{G}\mathbf{L}_s\mathbf{G}^\top\right). \tag{13}$$

Removing constant terms unrelated to $\mathbf{G}$, Eq. (13) is simplified to the following form,

$$\min_{\mathbf{G}} \operatorname{Tr}\left(\mathbf{G}\left(\frac{\lambda}{\beta}\mathbf{L}_s + v\mathbf{I}\right)\mathbf{G}^\top - 2\sum_{p=1}^{v} {\mathbf{W}^{(p)}}^\top \mathbf{Q}^{(p)}\mathbf{G}^\top \right). \tag{14}$$

Taking the derivative and making it equal to 0 yields:

$$\mathbf{G} = \sum_{p=1}^{v} {\mathbf{W}^{(p)}}^\top \mathbf{Q}^{(p)} \left(\frac{\lambda}{\beta}\mathbf{L}_s + v\mathbf{I}\right)^{-1}. \tag{15}$$

The entire procedure for solving Eq. (6) is outlined in Algorithm 1.

Complexity Analysis

Benefiting from the utilization of efficient computational techniques, the proposed SCSL exhibits a commendably low computational complexity. For instance, in solving the quadratic programming problem of Eq. (8), a closed-form solution is derived. Furthermore, to address the orthogonal constraint issue in Eq. (12), the algorithm employs the efficient Generalized Power Iteration Method (GPI). Additionally, instead of matrix inversion, a linear equation system is solved to compute Eq. (15), thereby contributing to a reduction in computational load.

The SCSL algorithm primarily encompasses three optimization steps, namely, updating $\mathbf{S}$, updating $\mathbf{W}^{(p)}$, and updating $\mathbf{G}$. The computational cost of updating $\mathbf{S}$ involves $\mathcal{O}(knd + kv^2n^2)$ operations for calculating Eq. (9), where $d = \sum_{p=1}^{v} d_p$. During the update of $\mathbf{W}^{(p)}$, matrix multiplication demands $\mathcal{O}(n^2d + (n+k)d^2)$ operations, while invoking the GPI algorithm for solving Eq. (12) accounts for $\mathcal{O}(kd^2 + k^2d)$ operations. When updating $\mathbf{G}$, matrix multiplication requires $\mathcal{O}(knd)$ operations, and solving the linear equation system necessitates $\mathcal{O}(kn^2)$ operations.

In summary, the computational complexity of the proposed SCSL algorithm is characterized by $\mathcal{O}(knd+kv^2n^2+n^2d+(n+k)d^2+k^2d)$. Notably, within the context of this algorithm, it holds that $k \ll n, v \ll n, d \ll n$, rendering the complexity quadratic for the number of samples.

Algorithm 1: SCSL

Require: Incomplete dataset $\{\mathbf{X}^{(p)}\}_{p=1}^{v}$, missing index matrix $\{\mathbf{H}^{(p)}\}_{p=1}^{v}$, normalized weight $\boldsymbol{\omega}$, parameters β, λ and the number of cluster k.

1: Initialize $\{\mathbf{W}^{(p)}\}_{p=1}^{v}$ and $\mathbf{G}$.
2: **repeat**
3: Update $\mathbf{S}$ by solving Eq. (7);
4: Update $\{\mathbf{W}^{(p)}\}_{p=1}^{v}$ by solving Eq. (11);
5: Update $\mathbf{G}$ by solving Eq. (13);
6: **until** converged.

Ensure: Perform spectral clustering on $\mathbf{S}$ to obtain cluster labels.

Dataset	Size	#Classes	#Views	#Features
MSRCV	210	7	6	256/512/1302
ORL	400	40	3	3304/4096/6750
ProteinFold	694	27	12	694/. . . /694
Wiki	2866	10	2	10/128
CCV	6773	20	3	20/20/20
SUNRGBD	10335	45	2	4096/4096

Table 2: Datasets used in our experiments.

Experiments

Experimental Settings

Datasets We evaluate the effectiveness of the proposed algorithm using six widely used datasets: MSRCV, ORL, ProteinFold, Wiki, CCV, and SUNRGBD. Detailed information about these datasets can be found in Table 2. To create incomplete versions of these datasets, we randomly remove samples from each view. Following the approach outlined in (Li et al. 2022), we ensure that each sample is present in at least one view. To assess the algorithm's performance under varying degrees of incompleteness, we generate incomplete datasets at intervals of 0.1, ranging from 0.1 to 0.9.

Datasets	MSRCV	ORL	ProteinFold	Wiki	CCV	SUNRGBD
			ACC			
BSV	33.36±2.45	24.32±0.89	22.25±0.53	44.75±1.26	11.86±0.10	6.14±0.08
MIC	47.19±3.12	37.56±1.66	15.99±0.78	40.50±1.74	12.85±0.67	14.61±0.54
MKKM-IK	61.51±1.77	59.80±2.44	26.03±1.06	45.23±0.22	14.72±0.33	11.35±0.31
DAIMC	66.04±7.40	68.03±2.32	28.65±1.65	46.07±1.11	15.17±0.73	**17.03±0.65**
UEAF	55.47±4.67	60.25±2.50	28.72±1.53	**47.75±0.05**	**17.08±0.32**	15.35±0.41
MKKM-MKC	57.45±4.03	64.95±2.62	17.99±0.83	46.21±0.06	15.76±0.15	16.81±0.49
V^3H	71.11±5.42	67.03±1.45	17.33±0.48	32.57±0.46	11.98±0.35	-
AGC-IMC	**74.72±0.18**	63.97±1.21	**31.13±0.93**	20.43±0.00	10.68±0.00	-
HCP-IMSC	63.38±0.19	**68.40±1.29**	18.92±0.88	30.64±0.05	11.31±0.08	10.70±0.00
Proposed	**79.21±0.78**	**71.55±2.20**	**32.62±1.06**	**48.19±0.19**	**17.70±0.27**	**17.61±0.40**
			NMI			
BSV	24.03±2.45	48.49±0.90	27.60±0.59	40.28±1.04	4.85±0.21	3.27±0.08
MIC	35.96±2.62	56.44±1.00	16.64±1.02	29.39±1.20	7.80±0.84	21.27±0.35
MKKM-IK	49.67±1.57	75.95±1.33	33.70±0.84	38.51±0.11	10.44±0.19	15.27±0.25
DAIMC	58.58±5.05	**82.89±1.06**	37.76±1.08	32.31±0.83	10.51±0.72	21.53±0.43
UEAF	46.31±2.81	76.16±1.25	38.18±0.88	**43.68±0.03**	**14.28±0.19**	**21.72±0.22**
MKKM-MKC	45.94±3.01	79.76±1.41	24.88±0.84	**38.01±0.06**	11.94±0.15	20.48±0.28
V^3H	64.84±3.25	81.05±0.61	22.75±0.53	19.03±0.21	10.53±0.25	-
AGC-IMC	**68.64±0.05**	78.36±0.39	**38.42±0.56**	12.20±0.00	1.31±0.00	-
HCP-IMSC	62.23±0.19	66.10±1.53	21.79±0.91	31.37±0.14	11.31±0.11	19.88±0.40
Proposed	**68.65±0.22**	**84.71±0.95**	**40.94±0.61**	36.33±0.15	**13.81±0.11**	**24.03±0.18**
			Purity			
BSV	36.15±2.25	26.80±0.92	25.89±0.60	46.88±1.18	12.76±0.14	13.06±0.16
MIC	49.08±2.79	40.81±1.40	19.78±0.84	44.00±1.22	16.51±0.61	32.36±0.59
MKKM-IK	62.31±1.57	62.79±2.11	30.91±1.04	51.02±0.12	18.66±0.27	27.06±0.45
DAIMC	68.44±6.10	**71.82±1.79**	34.99±1.54	49.62±0.85	18.77±0.59	**34.89±0.59**
UEAF	57.37±3.99	63.90±1.90	35.47±1.16	**51.13±0.04**	**19.60±0.38**	33.37±0.48
MKKM-MKC	58.79±3.47	67.68±2.34	22.73±0.87	47.96±0.07	19.42±0.20	32.92±0.50
V^3H	73.66±4.05	70.22±1.09	22.24±0.51	38.43±0.37	15.30±0.31	-
AGC-IMC	**76.47±0.16**	66.3±0.93	**36.23±0.79**	22.19±0.00	10.81±0.00	-
HCP-IMSC	64.73±0.19	70.88±1.17	16.82±0.72	30.48±0.05	11.00±0.09	7.32±0.38
Proposed	**79.32±0.54**	**74.81±1.76**	**38.54±0.84**	**51.18±0.19**	**21.72±0.20**	**35.93±0.31**

Table 3: ACC, NMI and Purity comparison of different IMVC algorithms on six partial datasets. '-' means out of CPU memory.

Metrics	Method	Datasets					
		MSRCV	ORL	ProteinFold	Wiki	CCV	SUNRGBD
ACC	w/	**79.21±0.78**	**71.55±2.20**	**32.62±1.06**	**48.19±0.19**	**17.70±0.27**	**17.61±0.40**
	w/o	66.43±56.21	70.38±83.62	29.18±37.54	38.92±29.35	14.65±9.77	15.89±20.42

Table 4: Experimental results of SCSL with and without the Laplacian regularization term. Due to space limitations, the results of the other two metrics are provided in the appendix.

Compared Methods In addition to the proposed SCSL algorithm, we compared it against eight state-of-the-art incomplete multi-view clustering methods. These methods include BSV (Ng, Jordan, and Weiss 2002), MIC (Shao, He, and Philip 2015), MKKM-IK (Liu et al. 2017), DAIMC (Hu and Chen 2018), UEAF (Wen et al. 2019), MKKM-MKC (Liu et al. 2020b), V^3H (Fang et al. 2020), AGC-IMC (Wen et al. 2021a), HCP-IMSC (Li et al. 2022).

For all the aforementioned algorithms, we configured their parameters within their recommended ranges. In our proposed method, we search β in $[0.001, 1, 10]$ and λ in $[0.001, 0.1, 1]$.

Evaluation To ensure robustness, we repeated each experiment 20 times and calculated the average performance along with the standard deviation. To evaluate the clustering performance, we utilized three commonly used metrics: accuracy (ACC), normalized mutual information (NMI), and Purity. All experiments were conducted on a desktop computer equipped with an Intel Core i9-10900X CPU, 64GB of RAM, and MATLAB 2020b (64-bit).

Experimental Results

Table 3 displays the clustering outcomes for the six benchmark datasets. Furthermore, Fig. 2 offers a comparison of

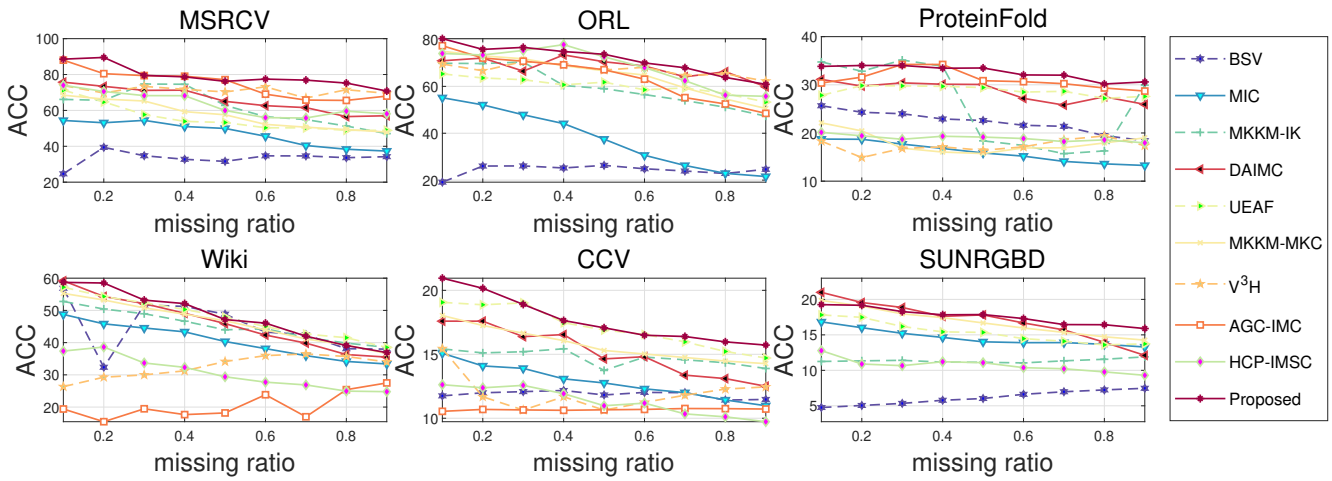


Figure 2: ACC metrics for clustering results on benchmark datasets with different missing rates. Due to space limitations, results for other metrics are provided in the appendix.

accuracy (ACC) across all methods at different missing rates. From the results, we draw the following conclusions:

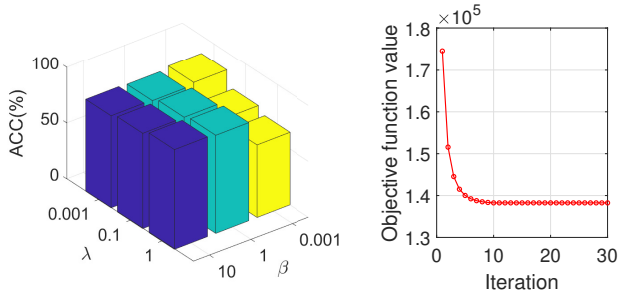
1. Our proposed SCSL algorithm generally outperforms existing IMVC methods on most datasets. Notably, the recently introduced HCP-IMSC method displays superior performance in incomplete datasets. In terms of ACC, SCSL surpasses the second-best method on the MSRCV, ORL, ProteinFold, Wiki, CCV, and SUNRGBD datasets by margins of 6.01%, 4.61%, 4.79%, 0.92%, 3.63%, and 3.41%, respectively.
2. When compared to traditional similarity-based IMVC methods in recent years (V^3H, AGC-IMC, HCP-IMSC), our approach generally achieves superior performance. This demonstrates the superiority of our proposed SCSL in constructing a comprehensive similarity by simultaneously utilizing the sample-level cross-view relationships.
3. As depicted in Fig. 2, most IMVC methods exhibit greater performance fluctuations with increasing missing rates, while SCSL maintains more stability. This trend suggests that the utilization of cross-view similarity compensates for the absent information from different views.

Ablation Study

To demonstrate the effectiveness of the Laplacian regularization term, we conducted experiments by removing this term in SCSL. As shown in Table 4, the addition of the Laplacian regularization term significantly improves the clustering performance of SCSL, validating its effectiveness in enhancing the quality of the similarity matrix.

Sensitivity and Convergence Analysis

We examined the sensitivity of SCSL to the parameters β and λ by investigating how our performance varies with their different value. As depicted in Fig. 3(a), the proposed method is not significantly affected by λ when β is large, while it is less influenced by β when λ is small.



(a) Study for β and λ (b) Convergence study

Figure 3: Sensitivity analysis of β and λ and convergence analysis of our method on MSRCV datasets. Convergence studies on other datasets are given in appendixs.

We performed a series of experiments to demonstrate the convergence behavior of SCSL. Fig. 3(b) illustrates that the objective value of our algorithm consistently decreases with each iteration, which provides clear evidence of the convergence of our proposed algorithm.

Conclusion

In this paper, we propose a novel incomplete multi-view clustering method termed as Sample-level Cross-view Similarity Learning (SCSL). Different from existing methods, we construct a cross-view similarity from every sample-pair among all views. Meanwhile, the consistency representation learning module contributes to improving the reliability of the projection and enhancing the quality of similarities through the inclusion of a graph Laplacian regularization. The proposed SCSL has been extensively evaluated on six benchmark datasets, showcasing its effectiveness in addressing the incomplete multi-view clustering problem when compared to state-of-the-art approaches.

Acknowledgements

This work was supported by the National Key R&D Program of China (project no. 2022ZD0209103), the National Natural Science Foundation of China (project no. 62325604, 62276271) and the Hunan Provincial Natural Science Foundation of China (project no. 2021JJ30779).

References

Deng, S.; Wen, J.; Liu, C.; Yan, K.; Xu, G.; and Xu, Y. 2023. Projective incomplete multi-view clustering. *IEEE Transactions on Neural Networks and Learning Systems*.

Deng, Z.; Liu, R.; Xu, P.; Choi, K.-S.; Zhang, W.; Tian, X.; Zhang, T.; Liang, L.; Qin, B.; and Wang, S. 2020. Multi-view Clustering with the Cooperation of Visible and Hidden Views. *IEEE Transactions on Knowledge and Data Engineering*.

Dong, Z.; Wang, S.; Jin, J.; Liu, X.; and Zhu, E. 2023. Cross-view Topology Based Consistent and Complementary Information for Deep Multi-view Clustering. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 19440–19451.

Fang, X.; Hu, Y.; Zhou, P.; and Wu, D. O. 2020. V ^{3} H: View Variation and View Heredity for Incomplete Multiview Clustering. *IEEE Transactions on Artificial Intelligence*.

Guo, J.; and Ye, J. 2019. Anchors bring ease: An embarrassingly simple approach to partial multi-view clustering. In *Proc. of AAAI*.

He, W.; Zhang, Z.; Chen, Y.; and Wen, J. 2023. Structured anchor-inferred graph learning for universal incomplete multi-view clustering. *World Wide Web*, 26(1): 375–399.

Hu, M.; and Chen, S. 2018. Doubly Aligned Incomplete Multi-view Clustering. In *Proc. of IJCAI*.

Jiang, G.; Peng, J.; Wang, H.; Mi, Z.; and Fu, X. 2022. Tensorial multi-view clustering via low-rank constrained high-order graph learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(8): 5307–5318.

Li, Z.; Tang, C.; Zheng, X.; Liu, X.; Zhang, W.; and Zhu, E. 2022. High-order correlation preserved incomplete multi-view subspace clustering. *IEEE Transactions on Image Processing*, 31: 2067–2080.

Liang, N.; Yang, Z.; and Xie, S. 2022. Incomplete multi-view clustering with sample-level auto-weighted graph fusion. *IEEE Transactions on Knowledge and Data Engineering*, 35(6): 6504–6511.

Lin, Y.; Gou, Y.; Liu, Z.; Li, B.; Lv, J.; and Peng, X. 2021. COMPLETER: Incomplete multi-view clustering via contrastive prediction. In *Proc. of CVPR*.

Liu, C.; Li, R.; Wu, S.; Che, H.; Jiang, D.; Yu, Z.; and Wong, H.-S. 2023. Self-guided partial graph propagation for incomplete multiview clustering. *IEEE Transactions on Neural Networks and Learning Systems*.

Liu, J.; Liu, X.; Zhang, Y.; Zhang, P.; Tu, W.; Wang, S.; Zhou, S.; Liang, W.; Wang, S.; and Yang, Y. 2021. Self-representation Subspace Clustering for Incomplete Multi-view Data. In *Proc. of ACM MM*.

Liu, S.; Liu, X.; Wang, S.; Niu, X.; and Zhu, E. 2022. Fast incomplete multi-view clustering with view-independent anchors. *IEEE Transactions on Neural Networks and Learning Systems*.

Liu, X.; Li, M.; Tang, C.; Xia, J.; Xiong, J.; Liu, L.; Kloft, M.; and Zhu, E. 2020a. Efficient and Effective Regularized Incomplete Multi-view Clustering. *IEEE transactions on pattern analysis and machine intelligence*.

Liu, X.; Li, M.; Wang, L.; Dou, Y.; Yin, J.; and Zhu, E. 2017. Multiple Kernel k-Means with Incomplete Kernels. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31.

Liu, X.; Zhu, X.; Li, M.; Wang, L.; Zhu, E.; Liu, T.; Kloft, M.; Shen, D.; Yin, J.; and Gao, W. 2020b. Multiple Kernel k k-Means with Incomplete Kernels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(05): 1191–1204.

Lv, Z.; Gao, Q.; Zhang, X.; Li, Q.; and Yang, M. 2022. View-consistency learning for incomplete multiview clustering. *IEEE Transactions on Image Processing*, 31: 4790–4802.

Ng, A. Y.; Jordan, M. I.; and Weiss, Y. 2002. On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*.

Nie, F.; Wang, X.; Jordan, M.; and Huang, H. 2016. The constrained laplacian rank algorithm for graph-based clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 30.

Nie, F.; Zhang, R.; and Li, X. 2017. A generalized power iteration method for solving quadratic problem on the stiefel manifold. *Science China Information Sciences*, 60: 1–10.

Shao, W.; He, L.; and Philip, S. Y. 2015. Multiple incomplete views clustering via weighted nonnegative matrix factorization with $\ell_{2,1}$ regularization. In *Joint European conference on machine learning and knowledge discovery in databases*.

Wan, X.; Liu, J.; Liang, W.; Liu, X.; Wen, Y.; and Zhu, E. 2022. Continual Multi-View Clustering. In *Proceedings of the 30th ACM International Conference on Multimedia*, MM '22, 3676–3684. New York, NY, USA: Association for Computing Machinery. ISBN 9781450392037.

Wang, H.; Zong, L.; Liu, B.; Yang, Y.; and Zhou, W. 2019. Spectral perturbation meets incomplete multi-view data. *arXiv preprint arXiv:1906.00098*.

Wang, Q.; Tao, Z.; Gao, Q.; and Jiao, L. 2022a. Multi-View Subspace Clustering via Structured Multi-Pathway Network. *IEEE Transactions on Neural Networks and Learning Systems*.

Wang, S.; Liu, X.; Liu, L.; Tu, W.; Zhu, X.; Liu, J.; Zhou, S.; and Zhu, E. 2022b. 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*, 9776–9785.

Wen, J.; Liu, C.; Deng, S.; Liu, Y.; Fei, L.; Yan, K.; and Xu, Y. 2023a. Deep Double Incomplete Multi-View Multi-Label Learning With Incomplete Labels and Missing Views. *IEEE Transactions on Neural Networks and Learning Systems*.

Wen, J.; Yan, K.; Zhang, Z.; Xu, Y.; Wang, J.; Fei, L.; and Zhang, B. 2021a. Adaptive Graph Completion Based Incomplete Multi-View Clustering. *IEEE Transactions on Multimedia*, 23: 2493–2504.

Wen, J.; Zhang, Z.; Fei, L.; Zhang, B.; Xu, Y.; Zhang, Z.; and Li, J. 2022. A survey on incomplete multiview clustering. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(2): 1136–1149.

Wen, J.; Zhang, Z.; Xu, Y.; Zhang, B.; Fei, L.; and Liu, H. 2019. Unified embedding alignment with missing views inferring for incomplete multi-view clustering. In *Proc. of AAAI*.

Wen, J.; Zhang, Z.; Xu, Y.; and Zhong, Z. 2018. Incomplete multi-view clustering via graph regularized matrix factorization. In *Proceedings of the European conference on computer vision (ECCV) workshops*, 0–0.

Wen, J.; Zhang, Z.; Zhang, Z.; Fei, L.; and Wang, M. 2020. Generalized incomplete multiview clustering with flexible locality structure diffusion. *IEEE transactions on cybernetics*.

Wen, J.; Zhang, Z.; Zhang, Z.; Zhu, L.; Fei, L.; Zhang, B.; and Xu, Y. 2021b. Unified tensor framework for incomplete multi-view clustering and missing-view inferring. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, 10273–10281.

Wen, Y.; Liu, S.; Wan, X.; Wang, S.; Liang, K.; Liu, X.; Yang, X.; and Zhang, P. 2023b. Efficient Multi-View Graph Clustering with Local and Global Structure Preservation. In *Proceedings of the 31st ACM International Conference on Multimedia*, MM '23, 3021–3030. New York, NY, USA: Association for Computing Machinery. ISBN 9798400701085.

Xia, D.; Yang, Y.; Yang, S.; and Li, T. 2023. Incomplete multi-view clustering via kernelized graph learning. *Information Sciences*, 625: 1–19.

Xu, C.; Guan, Z.; Zhao, W.; Wu, H.; Niu, Y.; and Ling, B. 2019. Adversarial Incomplete Multi-view Clustering. In *Proc. of IJCAI*.

Yang, X.; Jiaqi, J.; Wang, S.; Liang, K.; Liu, Y.; Wen, Y.; Liu, S.; Zhou, S.; Liu, X.; and Zhu, E. 2023. Dealmvc: Dual contrastive calibration for multi-view clustering. In *Proceedings of the 31st ACM International Conference on Multimedia*, 337–346.

Yang, Y.; Zhan, D.; Wu, Y.; Liu, Z.; Xiong, H.; and Jiang, Y. 2021. Semi-Supervised Multi-Modal Clustering and Classification with Incomplete Modalities. *IEEE Trans. Knowl. Data Eng.*, 33(2): 682–695.

Yin, J.; and Sun, S. 2022. Incomplete multi-view clustering with cosine similarity. *Pattern Recognition*, 123: 108371.

Yu, S.; Wang, S.; Wen, Y.; Wang, Z.; Luo, Z.; Zhu, E.; and Liu, X. 2023. How to Construct Corresponding Anchors for Incomplete Multiview Clustering. *IEEE Transactions on Circuits and Systems for Video Technology*.

Zhang, J.; Li, L.; Wang, S.; Liu, J.; Liu, Y.; Liu, X.; and Zhu, E. 2022. Multiple kernel clustering with dual noise minimization. In *Proceedings of the 30th ACM International Conference on Multimedia*, 3440–3450.