Preprint – IJCAI 2025: This is the accepted version made available for conference attendees.
Do not cite. The final version will appear in the IJCAI 2025 proceedings.

Robust Graph Contrastive Learning for Incomplete Multi-view Clustering

Deyin Zhuang[1], **Jian Dai**[2], **Xingfeng Li**[1], **Xi Wu**[1], **Yuan Sun**[3,4*], **Zhenwen Ren**[1*]
[1]Southwest University of Science and Technology, China
[2]Southwest Automation Research Institute, China
[3] College of Computer Science, Sichuan University, China
[4] National Key Laboratory of Fundamental Algorithms and Models for Engineering Numerical Simulation, Sichuan University, China
{zhuangdeyin163, daijian1000,xixi010129, sunyuan_work}@163.com,
{lixingfeng, rzw}@njust.edu.cn

Abstract

In recent years, multi-view clustering (MVC) has become a promising approach for analyzing heterogeneous multi-source data. However, during the collection of multi-view data, factors such as environmental interference or sensor failure often lead to the loss of view sample data, resulting in incomplete multi-view clustering (IMVC). Graph contrastive IMVC has demonstrated promising performance as an effective solution, which typically utilizes in-graph instances as positive pairs and out-of-graph instances as negative pairs. However, the construction of positive and negative pairs in this paradigm inevitably leads to graph noise Correspondence (GNC). To this end, we propose a new IMVC framework, namely robust graph contrastive learning (RGCL). Specifically, RGCL first completes the missing data by using a multi-view consistency transfer relationship graph. Then, to mitigate the impact of false negative pairs from graph contrastive, we propose noise-robust graph contrastive learning to mine intra-view consistency accurately. Finally, we present cross-view graph-level alignment to fully exploit the complementary information across different views. Experimental results on the six multi-view datasets demonstrate that our RGCL exhibits superiority and effectiveness compared with 9 state-of-the-art IMVC methods. The source code is available at https://github.com/DYZ163/RGCL.git.

1 Introduction

With the continuous development of data acquisition methods, multi-view data [Zhu *et al.*, 2025; Liang *et al.*, 2024; Xu *et al.*, 2025] could be collected by various devices, including cameras, infrared sensors, and audio sensors. As an unsupervised analysis technique, multi-view clustering (MVC) [Sun *et al.*, 2024; Li *et al.*, 2023b] aims to comprehensively analyze the cluster characteristics by leveraging the complementarity and consistency between the different views of the same instance. Existing MVC methods primarily depend on the assumption of the completeness of multi-view data. However, due to various uncontrollable factors during data collection, transmission, or storage, partial data could be missing, thereby leading to incomplete multi-view problems. This could disrupt the consistency between different views and create an imbalance in the available data information, which makes incomplete multi-view clustering (IMVC) more challenging.

To handle incomplete multi-view data, a variety of IMVC methods [Li *et al.*, 2025; Li *et al.*, 2023c] have been proposed, which could be roughly divided into two principal categories, i.e., shallow IMVC methods and deep IMVC methods. Thanks to the powerful feature representation capabilities of deep learning, deep IMVC methods [Xue *et al.*, 2021; Yuan *et al.*, 2025] can excavate more discriminative representations to uncover inter-cluster relationships in incomplete multi-view data, thereby significantly improving clustering performance. Therefore, in recent years, deep IMVC [Liu *et al.*, 2024] has attracted widespread attention from scholars.

To effectively explore the consistency information in incomplete multi-view data, some IMVC methods [Chao *et al.*, 2024; Tang and Liu, 2022] identify the cross-view neighbors through the observed samples and utilize them to estimate the missing data. However, these methods typically project all views into a common space to learn inter-view consistency, which unconsciously neglects view-specific information. To fully utilize the information from multiple views, some studies employ predictors [Lin *et al.*, 2022] and generators to capture view-specific information. Nevertheless, the generative paradigm is inherently focused on identifying a unified representation for the entire multi-view dataset, which may result in the loss of semantic information that is critical for clustering. To enhance the consistency and specificity of incomplete view data, many contrastive learning-based IMVC methods [Yuan *et al.*, 2024] have been proposed. For instance, COMPLETER [Lin *et al.*, 2021] adopts contrastive learning to maximize the mutual information between different views and proposes dual prediction to recover the missing data by minimizing conditional entropy. To further excavate consistency and achieve data restorability, MRL_CAL [Wang

*Corresponding Authors

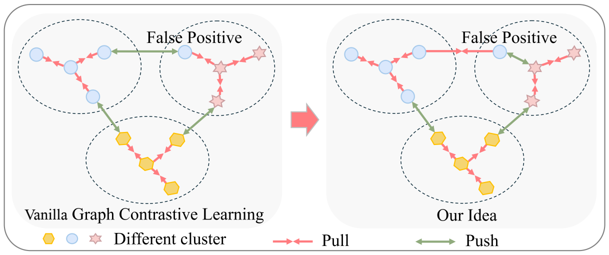


Figure 1: The motivation and key idea of our RGCL.

et al., 2024b] proposes adversarial learning to restore the missing data and utilizes contrastive learning to enhance the consistency of representations. Different from feature-level contrastive learning, AGCL [Wang *et al.*, 2022] constructs positive and negative sample pairs based on whether the sample is in the relationship graph, thereby achieving graph contrastive learning. To achieve cluster-level alignment, CGCN [Wang *et al.*, 2024c] proposes cross-view graph contrastive learning to enhance the discrimination of the learned representations.

Although existing graph contrastive IMVC methods have obtained delightful clustering performance, their superiority is heavily dependent on the assumption of an ideal graph structure. However, in practice, learning great graph relations is very challenging. In Fig.1, due to the inherent one-to-many graph contrastive characteristic and the strict distinction between positive and negative sample pairs, graph contrastive MVC inevitably mistakes semantically similar positive pairs as negative pairs, thus causing Graph Noisy Correspondences (GNC). In brief, GNC represents the false negative problem (FNP) caused by the noise presented in negative pairs.

To mitigate the aforementioned GNC problems, we propose a novel IMVC framework, namely **Robust Graph Contrastive Learning (RGCL)**. As shown in Fig.2, we first construct relation graphs by the similarities of nearest neighbors to search for similar samples. Further, based on the basis of semantic consistency, we adopt a graph transfer module to impute the missing samples of incomplete multi-view data. To alleviate the adverse impact of false negative pairs from graph contrastive, we design a target distribution sharpening strategy and propose noise-robust graph contrastive learning to explore intra-view consistency accurately. Finally, we present cross-view graph-level alignment to extract complementary information between different views. In general, the main contributions of this paper are summarized as follows:

- To overcome the negative influence caused by the false negative pairs, we propose a novel IMVC framework, namely robust graph contrastive learning (RGCL). To the best of our knowledge, this could be the first time to reveal and study a persistent more practical problem in graph contrastive MVC, dubbed graph noisy correspondence (GNC).
- We propose a novel noise-robust graph contrastive learning to prevent the interference of false negatives, thereby embracing the robustness of graph contrastive MVC against GNC. Moreover, we propose cross-view graph-level alignment to enhance cross-view consistency and complementarity.
- Extensive experiments on six popular multi-view datasets show the effectiveness and robustness of our RGCL method over state-of-the-art IMVC competitors.

2 Related Work

2.1 Incomplete Multi-view Clustering

In recent years, researchers have proposed numerous incomplete multi-view clustering (IMVC) methods. These methods can be classified into traditional IMVC and deep incomplete multi-view clustering (DIMVC). The traditional IMVC methods can be categorized into four categories: matrix factorization-based methods [Wen *et al.*, 2024], subspace-based methods, graph-based methods and kernel learning-based methods. Owing to the superior representation capabilities of deep neural networks for highly heterogeneous data [Liang *et al.*, 2025; Sun *et al.*, 2023], DIMVC has progressively emerged as a predominant research direction. DIMVC methods can be classified into three categories according to the structure of the deep learning model used: (1) Deep Neural Network-based methods (DNN), Zhang et al. [Zhang *et al.*, 2020] employ DNN to cluster learning, thus facilitating the discovery of the latent structure and information present within the data. (2)Generative Adversarial Networks-based (GAN) methods, Wang et al. [Wang *et al.*, 2023] used a GAN to fill in incomplete data and learn the consistency of multi-view data through bi-contrastive learning. (3)Autoencoder-based methods, URRL-IMVC [Teng *et al.*, 2024] employs autoencoders to extract features from complete data while utilizing inter-view correlations to impute missing data, thereby enabling the learning of consistent and clustering-friendly representations.

2.2 Contrastive Learning

Contrastive learning [Liang *et al.*, 2023], due to its powerful learning capabilities in areas such as image classification and deep clustering within computer vision, has become one of the most popular research topics in unsupervised learning. The core idea is to optimize the feature space by maximizing the similarity of positive sample pairs and minimizing the similarity of negative sample pairs. In DIMVC, contrastive learning effectively increases the compactness of samples within clusters and the separation between clusters. CC [Li *et al.*, 2021] introduces a novel online contrastive clustering method that leverages data augmentation to generate positive and negative sample pairs, enabling the alignment of instance features and the separation of clusters through contrastive learning. MCMVC [Geng *et al.*, 2022] and Dealmvc [Yang *et al.*, 2023] adopt a dual contrastive learning approach, employing instance-level and class-level contrastive learning to ensure consistency across views. However, traditional contrastive learning fails to fully exploit the local structure in incomplete multi-view data. To solve this problem, UGCF [Wang *et al.*, 2024a] proposed a unified graph contrastive learning framework, which recovers the missing data through the relationship saved in the view, and then performs graph

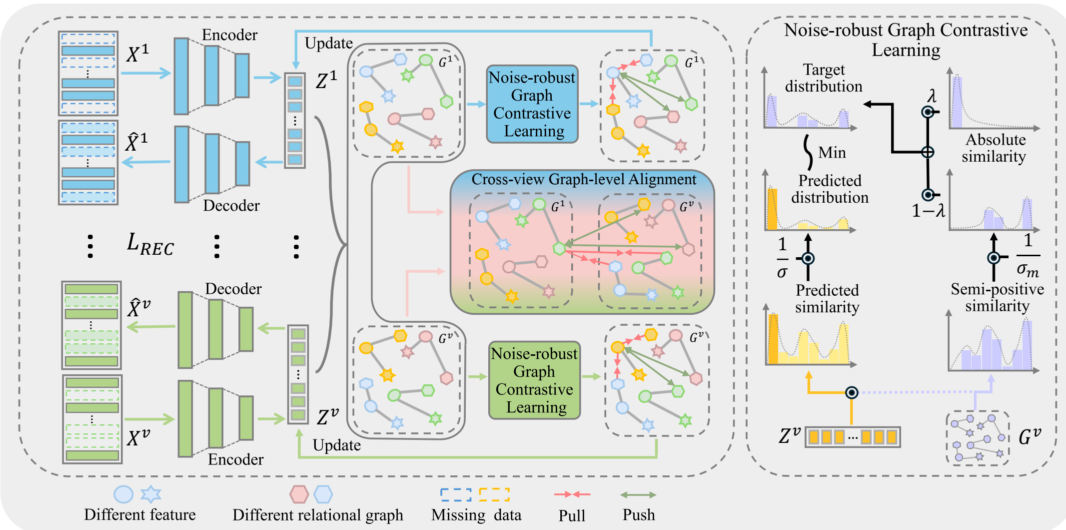


Figure 2: The framework of the proposed RGCL. First, RGCL employs the autoencoder to learn the specific embedding representation $\boldsymbol{Z}^v$ for each view $\boldsymbol{X}^v$. Further, the complete relation graph $\boldsymbol{G}^v$ is constructed based on $\boldsymbol{Z}^v$. Then, we propose noise-robust graph contrastive learning to reduce the influence of false negative pairs. Finally, we present a graph-level alignment strategy to fully use the complementary information between different views.

contrastive learning on the unified graph to improve the discrimination of features. Unlike existing studies, we reveal the graph-noise correspondence problem in IMVC. Further, we propose a robust graph contrastive learning framework to mitigate the impact of false negative pairs.

3 Method

3.1 Notations

For the IMVC task, our goal is to recover the missing data from different views to group the given incomplete dataset $\boldsymbol{X} = \{\boldsymbol{X}^1, \dots, \boldsymbol{X}^V\}$ with V views into different C clusters. Formally, we denote $\boldsymbol{X}^v = \{\boldsymbol{x}_1^v, \dots, \boldsymbol{x}_N^v\} \in \mathbb{R}^{N \times d_v}$ as the N sample data from the v-th view, where d_v and $\boldsymbol{x}_i^v$ represent the feature dimension and i-th sample from the v-th view, respectively. Note here, we use the mask matrix $\boldsymbol{M}^v = \{M_1^v, \dots, M_N^v\} \in \mathbb{R}^N$ to represent the available and missing case of N samples from the v-th view. Specifically, $M_i^v = 1$ indicates that there exists the i-th sample from the v-th view, otherwise, the corresponding sample is missing.

To achieve IMVC, we first construct a relationship graph for multi-view data to represent the similarities between the samples. However, since some views have missing data, we cannot construct a complete relationship graph by calculating the distance between the two views. To this end, according to [Wang *et al.*, 2022], we adopt the cross-view relationship graph transfer scheme to fill in the missing data, thereby constructing a complete relationship graph. As shown in Fig.3, for the missing data x_i^2 from the 2-th view, we can fill in the missing data by the relational graph of the corresponding sample x_i^1 from the 1-th view. Specifically, we first a K-nearest relational graph for $\boldsymbol{x}_i^1$, i.e., $\boldsymbol{G}_i^1 = \{\boldsymbol{x}_{i^1}^1, \dots, \boldsymbol{x}_{i^K}^1\}$, where K is the number of nearest neighbors. Then, we can utilize $\boldsymbol{G}_i^1$ to obtain the transfer relational graph $\boldsymbol{G}_i^2$. Since some neighbor samples of $\tilde{\boldsymbol{x}}_i^2$ in $\boldsymbol{G}_i^2$ are missing in the second view, we remove these missing neighbor samples to obtain the most accurate transition relationship graph, $\boldsymbol{G}_i^2 = \{\boldsymbol{x}_{i^1}^1, \dots, \boldsymbol{x}_{i^{K'}}^1\}$. Finally, we can calculate the mean of samples in the transferred relational graph $\boldsymbol{G}_i^2 = \{\boldsymbol{x}_{i^1}^1, \dots, \boldsymbol{x}_{i^{K'}}^1\}$ to complete the missing data $\boldsymbol{x}_i^2$. Mathematically, when $\boldsymbol{M}_i^v = 0$, the missing sample could be formalized as follows:

$$\tilde{\boldsymbol{x}}_i^v = \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{x}_{i^k}^v, \tag{1}$$

where $\boldsymbol{x}_{i^k}^v$ is the k-th sample in the transferred relation graph $\boldsymbol{G}_i^v$. $\tilde{\boldsymbol{x}}_i^v$ is the filled sample for the missing data $\boldsymbol{x}_i^v$.

3.2 Overview of RGCL

In this paper, we propose a new graph contrastive IMVC framework, namely RGCL. As illustrated in Fig. 2, RGCL first constructs the relation graph $\boldsymbol{G}^v$ and further uses graph transfer to complete the missing samples. Then, an autoencoder is used to encode the multi-view data $\boldsymbol{X}^v$ into the embedded representation $\boldsymbol{Z}^v$, which is reconstructed into the

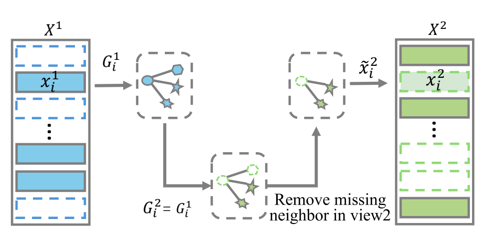


Figure 3: This is the process of completing data in incomplete multi-view datasets using relationship graphs.

negative pairs, we design the target distribution sharpening to achieve noise-robust graph contrastive learning and cross-view graph-level alignment, thereby endowing the robustness against GNC. In summary, the total loss function of our RGCL could be expressed as follows:

$$\mathcal{L} = \mathcal{L}_{REC} + \alpha\mathcal{L}_{NGC} + \beta\mathcal{L}_{CGA} + \mathcal{L}_{CLU}, \tag{2}$$

where L_{REC}, L_{NGC}, and L_{CGA} are multi-view reconstruction loss, noise-robust graph contrastive loss, and cross-view graph-level alignment loss, respectively. α and β are two hyper-parameters.

3.3 Multi-view Reconstruction

To obtain the view-specific representations, we employ the widely used multi-view reconstruction loss to learn more cluster-friendly representations from multi-view data. Specifically, after the multi-view completion, we first use the specific encoders $E^v_{\theta^v}(.)$ to extract the embedding representations $\boldsymbol{z}^v_i$ for the v-th view, which is formulated as:

$$\boldsymbol{z}^v_i = E^v_{\theta^v}\left(\boldsymbol{x}^v_i\right), \tag{3}$$

where θ^v is the parameters of the encoders $E^v_{\theta^v}(.)$. Note here that $\boldsymbol{z}^v_i \in \mathbb{R}^C$, where C is the number of clusters. Then, we adopt the specific decoders to obtain the reconstruction samples $\hat{\boldsymbol{x}}^v_i$, i.e.,

$$\hat{\boldsymbol{x}}^v_i = D^v_{\eta^v}(\boldsymbol{z}^v_i) = D^v_{\eta^v}\left(E^v_{\theta^v}(\boldsymbol{x}^v_i)\right), \tag{4}$$

where η^v is the parameters of the decoders $D^v_{\eta^v}(.)$. Therefore, the total multi-view reconstruction loss can be written as:

$$\mathcal{L}_{REC} = \frac{1}{VN}\sum_{v=1}^{V}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{x}}^v_i - M^v_i\boldsymbol{x}^v_i - (1-M^v_i)\tilde{\boldsymbol{x}}^v_i\right\|_2^2. \tag{5}$$

In other words, when $M^v_i = 1$, the multi-view reconstruction focuses on minimizing the reconstruction error of $\boldsymbol{x}^v_i$. If $M^v_i = 0$, it minimizes the bias in reconstructing $\tilde{\boldsymbol{x}}^v_i$.

3.4 Noise-robust Graph Contrastive Learning

To mitigate the adverse effects caused by false negatives in graph contrastive IMVC, we propose noise-robust graph contrastive learning, which leverages a relational graph to sharpen the target distribution, thereby endowing the robustness of the noise-robust graph contrastive learning. Specifically, in the training stage, we first use the K-Nearest Neighbors to construct a relational graph $\boldsymbol{G}^v_i = \{\boldsymbol{z}^v_{i^1}, ..., \boldsymbol{z}^v_{i^K}\}$ for the learned embedding representations $\boldsymbol{z}^v_i$. In order to prevent semantically similar positive pairs from being mistaken for negative pairs, we propose the concept of semi-positive pairs. To be specific, for each representation, we consider the samples in the same relationship graph to be absolute positive pairs, while outside the relational graph but have high similarity to be semi-positive pairs. The samples with low similarity outside the relational graph are regarded as negative pairs.

Inspired by [Zheng *et al.*, 2021], we design a target distribution sharpening strategy to measure the similarity of semi-positive pairs, which could prevent the introduction of false negative pairs. The similarity distribution $\boldsymbol{s}^{vk}_i = [s^{vk}_{i0}, s^{vk}_{i1}, \ldots, s^{vk}_{iN}] \in \mathbb{R}^{N-1}$ between semi-positive samples can be represented as:

$$s^{vk}_{im} = \frac{\mathbb{1}_{i\neq m}\cdot\exp\left(\boldsymbol{z}^v_{i^k}\cdot\boldsymbol{z}^v_{m^k}/\theta\right)}{\sum_{j=1}^{N}\mathbb{1}_{i\neq j}\cdot\exp\left(\boldsymbol{z}^v_{i^k}\cdot\boldsymbol{z}^v_{j^k}/\theta\right)}, \tag{6}$$

where $\boldsymbol{z}^v_{i^k}$, $\boldsymbol{z}^v_{m^k}$, and $\boldsymbol{z}^v_{j^k}$ represent the embedding representations in the relational graphs of $\boldsymbol{z}^v_i$, $\boldsymbol{z}^v_m$, and $\boldsymbol{z}^v_j$, respectively. Note here that $\boldsymbol{z}^v_i$, $\boldsymbol{z}^v_m$, and $\boldsymbol{z}^v_j$ are in the same mini-batch. To maintain the dominant role of absolute positive pairs in noise-robust graph contrastive learning, we set the similarity of the absolute positive pairs to 1. Then, we joint the similarity distribution $\boldsymbol{s}^{vk}_i$ and the similarity of the absolute positive pairs to obtain the target similarity distribution $\boldsymbol{w}^{vk}_i = [w^{vk}_{i0}, w^{vk}_{i1}, \ldots, w^{vk}_{iN}] \in \mathbb{R}^N$, i.e.,

$$w^{vk}_{im} = \lambda\cdot\mathbb{1}_{i=m} + (1-\lambda)\cdot s^{vk}_{im}, \tag{7}$$

where $\mathbb{1}_{i=m}$ denotes the indicator function, i.e. $\mathbb{1}_{i=m} = 1$ when $i = m$, otherwise $\mathbb{1}_{i=m} = 0$. λ is the weighting factor and is fixed as 0.5 in our paper.

To bring absolute positive pairs and semi-positive pairs closer while separating negative pairs, we hope to minimize the divergence between the predicted similarity distribution $\boldsymbol{p}^{vk}_i$ and the target similarity distribution $\boldsymbol{w}^{vk}_i$. Specifically, we first construct the predicted similarity distribution $\boldsymbol{p}^{vk}_i$ by calculating the similarity between $\boldsymbol{z}^v_i$ and the relationship graph $\boldsymbol{z}^v_{m^k}$ within the same mini-batch. Thus, we can obtain the predicted similarity distribution as follows:

$$p^{vk}_{im} = \frac{\exp\left(\boldsymbol{z}^v_i\cdot\boldsymbol{z}^v_{m^k}/\sigma\right)}{\sum_{j=1}^{N}\exp\left(\boldsymbol{z}^v_i\cdot\boldsymbol{z}^v_{j^k}/\sigma\right)}. \tag{8}$$

The temperature parameter σ is used to adjust the predicted similarity distribution $\boldsymbol{p}^{vk}_i$, where $\sigma > \theta$ makes the constructed predicted similarity distribution smoother. Clearly, we can obtain $\boldsymbol{p}^{vk}_i = [p^{vk}_{i0}, p^{vk}_{i2}, \ldots, p^{vk}_{iN}] \in \mathbb{R}^N$.

In general, we could assume that samples situated within the same relational graph and those situated outside it with high similarity should be pulled closer together. That is to say, absolute positive pairs and semi-positive pairs should be pulled closer together. Conversely, samples situated outside the same relational graph with low similarity should be pushed apart. By treating samples with high similarity from different relational graphs as semi-positive pairs, L^k_i can effectively use the repulsion effect of negative pairs. At the same time, it avoids overly penalizing negative samples that are semantically similar to positive samples, thereby enhancing the discriminative power of the embedding representations. Therefore, the loss function L^k_i could be calculated by the cross-entropy between $\boldsymbol{w}^k_i$ and $\boldsymbol{p}^k_i$, i.e.,

$$\mathcal{L}^k_i = -\frac{1}{V}\sum_{v=1}^{V}\sum_{m=1}^{N} w^{vk}_{im}\log\left(p^{vk}_{im}\right). \tag{9}$$

Finally, the total noise-robust graph contrastive learning loss can be described as follows:

$$\mathcal{L}_{NGC} = \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=1}^{K}\mathcal{L}^k_i. \tag{10}$$

3.5 Cross-view Graph-level Alignment

To fully exploit the inter-view complementarity information, we propose cross-view graph-level distribution alignment. Specifically, the cross-view predicted distribution $c_i^{vhk} = [c_{i0}^{vhk}, c_{i1}^{vhk}, \dots, c_{iN}^{vhk}]$ can be expressed as follows:

$$c_{im}^{vhk} = \frac{\exp\left(\boldsymbol{z}_i^v \cdot \boldsymbol{z}_{m^k}^h / \sigma\right)}{\sum_{j=1}^{N} \exp\left(\boldsymbol{z}_i^v \cdot \boldsymbol{z}_{j^k}^h / \sigma\right)}, \tag{11}$$

where $\boldsymbol{z}_{m^k}^h$ and $\boldsymbol{z}_{j^k}^h$ represent the relationship graphs of $\boldsymbol{z}_m^h$ and $\boldsymbol{z}_j^h$, respectively.

According to Eq.7, the cross-view graph-level target distributions w_i^{hk} are obtained from different relational graphs. The predictive distribution of the cross-view $\boldsymbol{c}_i^{\nu hk}$ and the target distribution of the cross-view graph $\boldsymbol{w}_i^{hk}$ are aligned by cross-entropy to achieve the alignment of the cross-view graph-level distribution. The loss function L_i^{vk} is as follows:

$$\mathcal{L}_i^{vk} = -\frac{1}{V}\sum_{h=1}^{V}\sum_{m=1}^{N} \mathbb{1}_{v\neq h} \cdot w_{im}^{hk} \log\left(c_{im}^{vhk}\right). \tag{12}$$

By extending Eq.12 from z_i^ν to the entire dataset, the final cross-view graph-level alignment loss can be written as:

$$\mathcal{L}_{CGA} = \frac{1}{NVK}\sum_{i=1}^{N}\sum_{v=1}^{V}\sum_{k=1}^{K} \mathcal{L}_i^{vk}. \tag{13}$$

In summary, this cross-view graph-level alignment can effectively ensure the distribution of $\boldsymbol{c}_i^{\nu hk}$ aligns with $\boldsymbol{w}_i^{hk}$, thereby fully leveraging the complementary information between views.

3.6 Implementation

The overall training process of our proposed method mainly contains two stages, i.e., warm-up and fine-tuning. The warm-up stage is summarized in the supplementary material. Afterward, we fine-tune the network during the second-stage training using clustering loss. Specifically, to obtain clustering predictions, a parametric mapping is employed for each $\boldsymbol{z}_i^v$, thereby resulting in the generation of soft clustering assignments, represented by the variable q_{ij}^v. Here, μ_j^v denotes the centroid of the j-th cluster in the v-th view, while q_{ij}^v indicating the probability of the embedding representation $\boldsymbol{z}_i^v$ being assigned to the j-th cluster. Given that the embedding representations of different views for a given sample exhibit similarity following learning, a fusion representation, denoted by $\boldsymbol{z}_i^*$, is obtained by summing these embedding representations. Where μ_j^* denotes the center of the fusion representation $\boldsymbol{z}_i^*$ and q_{ij}^* denotes the probability that the fusion representation $\boldsymbol{z}_i^*$ is assigned to the j-th cluster. Then, the assignment probability with the highest probability value is selected as the high-confidence assignment probability q_{ij}:

$$\begin{aligned} q_{ij}^v &= \frac{(1+||z_i^v-\mu_j^v||^2)^{-1}}{\sum_{c=1}^{C}(1+||z_i^*-\mu_c||^2)^{-1}}, \\ q_{ij}^* &= \frac{(1+||z_i^*-\mu_j^*||^2)^{-1}}{\sum_{c=1}^{C}(1+||z_i^*-\mu_c^*||^2)^{-1}}, \\ q_{ij} &= \max\{q_{ij}^1, q_{ij}^2, \cdots, q_{ij}^v, q_{ij}^*\}, \end{aligned} \tag{14}$$

where μ_j^v and μ_j^* are initialized on $\boldsymbol{z}_i^v$ and $\boldsymbol{z}_i^*$ by K-means [Vassilvitskii and K-means+, 2006], respectively. $\boldsymbol{Q} = [q_{ij}^*]$ is the distribution of high-confidence assignment probabilities. P represents the high-confidence target distribution of Q, which could be computed by the following formula:

$$p_{ij} = \frac{q_{ij}^2 / \sum_{i=1}^{N} q_{ij}}{\sum_{c=1}^{C}(q_{ic}^2 / \sum_{i=1}^{N} q_{ic})}. \tag{15}$$

Finally, following [Xu *et al.*, 2022], we fine-tune the network by using a clustering loss based on KL divergence to obtain the final clustering results. By adjusting Q to make it as close as possible to P, high-confidence predictions are used to generate representations more suitable for clustering. In conclusion, we obtained a more discriminative fused representation z^*. The clustering loss can be expressed as follows:

$$\mathcal{L}_{CLU} = KL(P||Q) = \sum_{i=1}^{N}\sum_{j=1}^{C} p_{ij} \boldsymbol{log}\frac{p_{ij}}{q_{ij}}. \tag{16}$$

4 Experiments

4.1 Datasets

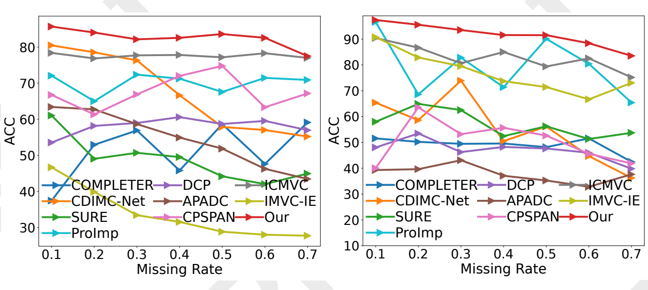
In this section, we evaluate the performance of the proposed method on the six multi-view datasets, including **HandWritten** [LeCun *et al.*, 1989], **COIL20** [Nene *et al.*, 1996], **BDGP** [Cai *et al.*, 2012], **LandUse-21** [Yang and Newsam, 2010], **ALOI-100** [Geusebroek *et al.*, 2005], and **AWA** [Romera-Paredes and Torr, 2015]. The details of all multi-view datasets we illustrate in the supplementary material.

4.2 Baselines and Metrics

To evaluate the effectiveness of our RGCL, we compare it with nine state-of-the-art IMVC methods, i.e., **CDIMC-Net** [Wen *et al.*, 2020], **COMPLETE** [Lin *et al.*, 2021], **SURE** [Yang *et al.*, 2022], **ProImp** [Li *et al.*, 2023a], **DCP** [Lin *et al.*, 2022], **APADC** [Xu *et al.*, 2023], **CPSPAN** [Jin *et al.*, 2023], **ICMVC** [Chao *et al.*, 2024], and **IMVC-IE** [Huang *et al.*, 2024]. In our experiments, we adopt three popular evaluation metrics (i.e., ACC, NMI, and ARI) to comprehensively assess the effectiveness of these methods.

4.3 Experimental Settings

In our experiments, we use full connected layers to construct the view-specific autoencoders. To be specific, the view-



(a) COIL20 (b) BDGP

Figure 4: The performance on two datasets with different missing rates.

		HandWritten			COIL20			ALOI-100			BDGP			LandUse-21			AWA		
	Metrics	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
0.1	CDIMC-Net(IJCAI'20)	74.30	70.80	59.17	80.48	88.57	77.45	35.37	59.76	12.49	65.40	49.60	34.18	22.61	31.87	10.75	50.35	63.85	36.66
	COMPLETER(CVPR'21)	71.55	77.18	63.24	37.43	58.45	29.21	21.85	57.39	14.23	51.55	41.91	13.98	21.88	29.29	10.96	25.63	45.88	17.70
	SURE(TPAMI'22)	73.85	62.22	53.43	61.04	76.45	52.23	33.71	71.34	21.70	58.00	35.97	27.65	24.52	28.48	10.93	34.22	52.61	26.93
	ProImp(IJCAI'23)	82.45	80.21	72.95	72.08	80.15	65.29	62.80	78.05	50.17	96.62	90.79	92.03	24.81	28.33	11.84	38.98	51.98	28.07
	DCP(PAMI'23)	72.70	76.51	56.98	53.54	72.18	42.22	23.18	58.60	13.27	48.06	45.51	18.71	25.11	31.29	12.03	27.18	47.00	17.79
	APADC(TIP'23)	81.23	78.81	71.38	63.40	73.57	52.54	36.32	64.61	22.01	39.31	22.31	6.84	17.34	24.33	5.34	34.05	48.61	22.81
	CPSPAN(CVPR'23)	87.15	79.59	74.47	66.74	78.30	59.84	54.85	74.80	41.41	90.44	77.78	77.08	21.05	30.25	10.65	47.29	60.23	37.37
	ICMVC(AAAI'24)	84.18	82.97	76.87	78.39	87.04	75.50	61.05	78.56	48.13	90.42	84.51	83.15	26.32	30.07	13.51	45.08	61.82	35.53
	IMVC-IE(ICASSP'24)	65.05	54.50	45.11	46.66	53.88	33.06	38.53	63.50	26.89	90.76	75.62	78.51	21.76	24.05	8.89	24.92	39.65	23.74
	RGCL(Our)	**92.30**	**85.71**	**83.88**	**84.51**	**90.83**	**82.63**	**71.41**	**82.24**	**59.38**	**97.40**	**91.37**	**93.64**	**28.24**	**34.31**	**13.80**	**59.92**	**68.87**	**49.82**
0.3	CDIMC-Net(IJCAI'20)	64.40	64.13	43.72	76.13	84.91	70.10	28.57	55.57	12.45	73.92	67.60	62.62	21.57	32.14	9.60	47.42	59.30	33.44
	COMPLETER(CVPR'21)	75.21	73.04	55.10	56.89	72.28	44.88	22.14	56.66	11.37	49.46	44.33	17.47	22.09	29.18	10.50	23.72	42.86	15.47
	SURE(TPAMI'22)	80.35	67.88	63.05	50.69	73.36	49.20	29.82	67.75	19.86	62.52	56.16	44.84	26.19	29.85	12.10	27.28	24.28	21.21
	ProImp(IJCAI'23)	81.60	79.31	71.19	72.36	79.12	59.20	60.53	76.73	48.67	82.84	71.61	67.19	24.86	28.38	11.84	37.66	50.83	25.34
	DCP(PAMI'23)	77.59	79.07	63.75	58.96	72.92	47.38	22.76	57.79	11.48	46.26	41.17	14.45	24.32	30.45	12.11	24.65	44.03	16.19
	APADC(TIP'23)	31.91	47.61	17.34	16.80	21.52	4.87	43.09	33.41	11.02	33.10	62.67	23.36	16.80	21.54	4.87	31.91	47.61	17.34
	CPSPAN(CVPR'23)	86.70	78.43	73.47	66.88	78.31	58.70	48.69	69.81	34.01	91.96	78.31	80.70	24.67	31.30	11.73	49.92	60.99	38.91
	ICMVC(AAAI'24)	81.95	79.36	72.25	77.71	86.15	74.66	61.98	78.40	47.30	80.67	72.97	69.48	26.14	29.11	12.24	48.20	62.56	38.28
	IMVC-IE(ICASSP'24)	46.98	46.98	40.36	33.40	37.51	14.26	38.06	61.83	25.76	79.52	53.30	55.88	19.76	19.20	6.45	24.85	39.88	22.09
	RGCL(Our)	**90.95**	**79.41**	**75.38**	**83.47**	**89.66**	**81.03**	**69.23**	**81.02**	**56.93**	**93.52**	**81.55**	**84.58**	**28.10**	**34.73**	**14.77**	**57.66**	**64.03**	**45.04**
0.5	CDIMC-Net(IJCAI'20)	50.90	52.17	33.39	57.86	71.80	46.28	24.17	52.43	8.85	56.16	33.25	24.44	20.61	30.69	9.56	41.44	55.29	27.75
	COMPLETER(CVPR'21)	66.60	68.59	42.95	58.64	74.75	49.31	18.69	52.89	10.43	48.17	38.97	15.01	20.97	25.44	9.48	23.85	43.68	15.03
	SURE(TPAMI'22)	80.10	69.03	63.69	44.17	69.15	43.62	28.94	67.29	20.04	56.24	33.25	25.86	24.00	27.50	10.64	28.91	42.11	19.57
	ProImp(IJCAI'23)	75.00	74.61	66.06	67.57	78.42	60.50	57.00	74.18	44.49	90.12	75.66	77.14	22.67	26.76	10.18	36.36	47.62	22.71
	DCP(PAMI'23)	71.43	75.40	57.12	58.68	73.57	50.06	21.50	56.38	8.96	47.64	39.97	13.96	23.25	28.81	9.48	24.14	43.15	14.60
	APADC(TIP'23)	70.35	68.96	57.00	51.85	65.67	41.30	30.53	60.96	21.92	35.26	23.51	4.28	17.39	22.31	5.61	34.54	47.55	19.26
	CPSPAN(CVPR'23)	79.95	73.94	65.99	74.79	81.86	66.72	51.48	69.62	37.90	76.48	61.55	55.73	24.57	31.13	9.73	52.81	**63.42**	**42.47**
	ICMVC(AAAI'24)	74.92	71.84	63.13	77.17	85.17	73.71	60.27	76.82	44.88	79.40	67.52	64.96	25.14	26.94	11.46	47.93	60.21	36.11
	IMVC-IE(ICASSP'24)	58.25	43.41	35.26	28.81	32.54	11.51	34.39	60.55	23.30	71.44	40.46	42.07	16.80	14.28	4.20	21.58	36.33	21.41
	RGCL(Our)	**86.50**	**79.10**	**74.46**	**82.77**	**88.54**	**79.88**	**70.23**	**80.10**	**56.91**	**91.52**	**76.74**	**80.14**	**27.76**	**33.63**	**13.94**	**53.99**	60.60	40.50
0.7	CDIMC-Net(IJCAI'20)	49.35	53.15	28.98	55.20	72.47	44.51	25.92	55.28	10.54	36.24	18.31	8.56	20.23	25.35	7.90	37.00	51.33	24.07
	COMPLETER(CVPR'21)	78.77	73.82	62.44	59.11	73.37	50.72	19.31	53.18	7.68	42.49	38.41	14.57	20.65	25.43	7.96	22.53	40.63	14.68
	SURE(TPAMI'22)	79.15	67.13	61.99	44.93	63.82	40.32	36.19	64.92	15.60	53.72	38.66	29.86	23.38	27.29	9.72	30.55	46.10	20.18
	ProImp(IJCAI'23)	78.40	74.01	67.16	70.89	78.25	62.26	55.84	73.98	44.02	65.44	54.36	44.54	23.29	26.00	9.79	34.75	47.34	23.15
	DCP(PAMI'23)	66.31	68.17	43.10	56.96	72.33	35.75	20.88	54.47	8.10	39.82	30.95	7.19	19.74	26.88	4.99	23.24	41.02	14.55
	APADC(TIP'23)	53.50	63.69	45.84	43.43	59.11	32.07	26.59	56.60	16.89	37.64	24.36	5.32	15.92	20.42	4.45	32.74	44.57	17.85
	CPSPAN(CVPR'23)	77.55	70.58	63.89	67.15	78.78	59.69	45.91	66.50	31.16	74.56	54.68	49.97	25.85	**32.08**	11.37	41.16	54.61	30.01
	ICMVC(AAAI'24)	72.38	69.77	59.52	77.01	84.41	73.13	50.50	71.12	38.75	75.18	58.98	55.75	21.64	25.63	8.20	42.25	55.78	31.75
	IMVC-IE(ICASSP'24)	67.65	53.98	45.87	27.70	26.32	10.01	34.32	60.08	22.80	73.04	46.35	46.60	19.38	19.00	6.17	19.88	29.88	18.31
	RGCL(Our)	**83.15**	**74.23**	**68.50**	**77.50**	**85.12**	**73.29**	**66.24**	**77.31**	**52.27**	**83.52**	**62.04**	**63.80**	**26.29**	31.83	**12.00**	**50.21**	**56.68**	**34.52**

Table 1: Performance comparison across six datasets with four distinct missing rates. The best results are highlighted in black.

specific encoder and decoder layers are configured with dimensions of ($0.8d_v$, $0.8d_v$, 1500, C) and (C, 1500, $0.8d_v$, $0.8d_v$, d_v), respectively. Note here, d_v and C represent the feature dimension from each view and the number of clustering categories, respectively. We set the temperature parameters to $\sigma = 0.1$ and $\theta = 0.05$. To evaluate the performance for incomplete multi-view data, we randomly set the instances with a certain ratio (i.e., [0.1, 0.3, 0.5, 0.7]) as the missing pairs. For all experiments, we employ a Linux platform equipped with an NVIDIA RTX 4090 GPU and 32GB of memory, using PyTorch version 2.3.0.

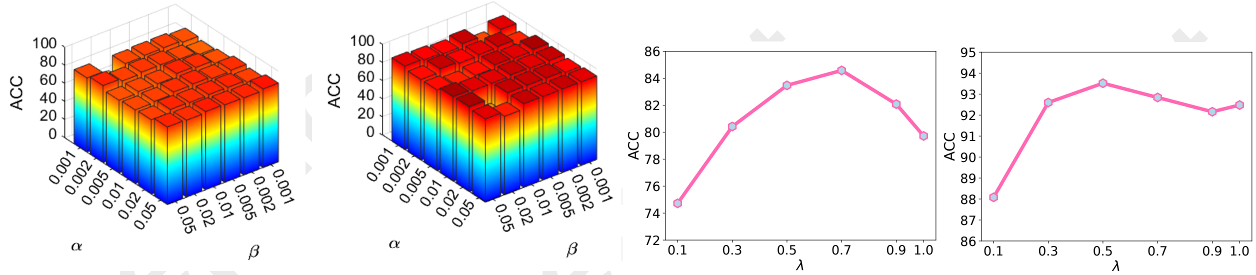
4.4 Experimental Analysis

The experimental results on the six datasets are shown in Tab.1. To visually illustrate the performance trends exhibited by each method as the missing rate varies, we plot the Fig.4. By analyzing the data in the above tables and figures, the following conclusions can be drawn:

Dataset	HandWritten		COIL20		BDGP		LandUse-21		ALOI-100		AWA	
MR	0.3	0.7	0.3	0.7	0.3	0.7	0.3	0.7	0.3	0.7	0.3	0.7
RGCL-1	83.25	79.00	78.06	69.37	80.48	63.60	21.81	22.95	38.21	38.22	49.70	42.73
RGCL-2	90.50	82.95	81.94	72.29	84.48	80.44	27.71	25.67	62.01	61.54	56.12	47.89
RGCL-3	89.30	82.85	79.37	74.24	88.84	69.40	27.29	24.24	48.40	50.93	54.60	46.93
RGCL-4	**90.95**	**83.15**	**83.47**	**77.50**	**93.52**	**83.52**	**28.10**	**26.29**	**69.23**	**66.24**	**57.66**	**50.21**

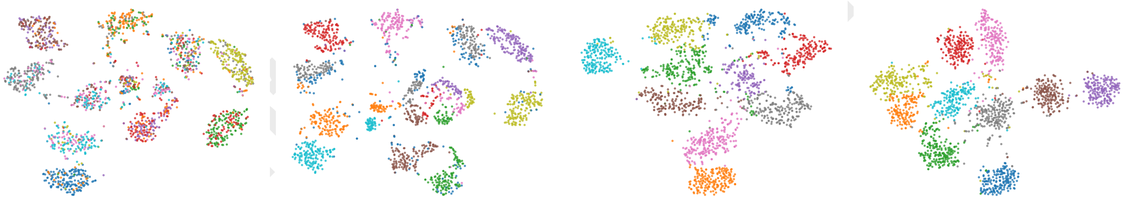
Table 2: Ablation studies on six datasets with different missing rates, where MR indicates the missing rates.

- In six datasets, our RGCL method outperforms nine other IMVC methods in most cases, demonstrating higher overall metrics. Specifically, on the ALOI-100 dataset with a 0.7 missing rate, RGCL achieves significant performance gains over the best baseline, ProImp. It improves ACC by 10.4%, NMI by 3.33%, and ARI by 8.25%.In contrast, the performance of other methods tends to decline significantly. This indicates that the RGCL is more accurate in representing incomplete multi-view data and identifying cluster relationships.
- Compared to several baselines based on contrastive learning, RGCL shows superior performance in clustering for all missing rates. On the BDGP dataset with a 0.1 missing rate, RGCL achieves a 45.85% improvement in ACC over COMPLETER and a 6.64% improvement over IMVC-IE. Moreover, as the missing rate increases in the incomplete multi-view dataset, the performance improvement of all evaluation metrics becomes more obvious. The results show that the proposed noise-robust graph contrastive learning effectively alleviates the impact of false negative pairs. Additionally, the cross-view graph-level alignment fully leverages the complementary information between views.
- On all datasets, RGCL shows lower variability in ACC, NMI, and ARI values as the missing rate increases, reflecting the stability of the model. In comparison to methods such as DCP and CPSPAN, RGCL demonstrates smoother values, indicating its broad applicability across various multi-view datasets and resilience to



(a) COIL20 (b) BDGP (c) COIL20 (d) BDGP

Figure 5: Parameter sensitivity analyses on the two datasets with 0.3 missing rate.



(a) APADC (b) ICMVC (c) CPSPAN (d) Our RGCL

Figure 6: Visualization results on HandWritten with 0.5 missing rate.

data fluctuations.

4.5 Ablation Study

To investigate the importance of each component in RGCL, we conduct an ablation study by isolating $\mathcal{L}_{\text{REC}}$, $\mathcal{L}_{\text{CLU}}$, $\mathcal{L}_{\text{NGC}}$, and $\mathcal{L}_{\text{CGA}}$ to verify their significance. As shown in Tab.2, four sets of experiments were conducted with different missing rates on six datasets. Here, RGCL-1 represents the use of components $\mathcal{L}_{\text{REC}}$ and $\mathcal{L}_{\text{CLU}}$, RGCL-2 represents the use of components $\mathcal{L}_{\text{REC}}$, $\mathcal{L}_{\text{CLU}}$, and $\mathcal{L}_{\text{NGC}}$, RGCL-3 represents the use of components $\mathcal{L}_{\text{REC}}$, $\mathcal{L}_{\text{CLU}}$, and $\mathcal{L}_{\text{CGA}}$, and RGCL-4 represents the use of all components. The results from these experiments indicate that $\mathcal{L}_{\text{REC}}$ and $\mathcal{L}_{\text{CLU}}$ plays a critical role in the autoencoder. Nevertheless, adding either $\mathcal{L}_{\text{NGC}}$ or $\mathcal{L}_{\text{CGA}}$ further improves performance, suggesting that both $\mathcal{L}_{\text{NGC}}$ and $\mathcal{L}_{\text{CGA}}$ support learning cluster-friendly representations. Optimal performance is achieved when all three losses are utilized. It is proved that $\mathcal{L}_{\text{NGC}}$ mitigates the adverse effects of false negatives, thereby preventing semantic information loss. Meanwhile, $\mathcal{L}_{\text{CGA}}$ maximizes the utilization of complementary information across views.

4.6 Parameter Sensitivity Analysis

In this section, we first analyze the effect of the two hyperparameters (i.e., α and β) in our loss on the whole model, we conduct parameter sensitivity analysis on the COIL-20 and BDGP datasets. Specifically, we set the missing rate to 0.5 and vary the parameter range from 0.001 to 0.05. As shown in Fig.5(a)-(b), excessively large or small parameter values adversely affect clustering performance.

Further, we analyze the effect of parameter λ on solving the false negative problem. In Fig.5(c)-(d), the ACC results of the parameter λ from 0.1 to 1 process are presented. The overall trend of ACC shows a gradual increase at first, followed by a subsequent decrease. The parameter λ can balance the relationship between positive sample pairs and semi-positive sample pairs. It shows that noise-robust graph contrastive learning can effectively mine the information beneficial to clustering in false negative samples, thereby alleviating the false negative problem. Thus, based on the analysis of the experimental results, we obtain the optimal value of λ,α and β, i.e. $\lambda = 0.5$, α=0.005 or 0.01, and β=0.005 or 0.01.

4.7 Visualization

To illustrate the clustering advantage of RGCL, we compare it with three state-of-the-art methods (APADC, ICMVC, and CPSPAN) on the HandWritten dataset with 0.5 missing rate. As shown in Fig.6, ICMVC shows large intra-cluster and small inter-cluster dispersion, while RGCL achieves more compact intra-cluster and more dispersed inter-cluster distributions, highlighting its superior clustering capability.

5 Conclusion

In this paper, we reveal and study a practical graph noisy correspondence (GNC) in the field of graph contrastive IMVC. To be specific, graph contrastive learning could lead to false negative pairs caused by the inherent one-to-many graph contrastive characteristic. To overcome this problem, we propose a novel robust graph contrastive learning framework (RGCL) for IMVC. Specifically, RGCL first completes the missing data through the multi-view consistency transition relationship graph. Afterward, we propose noise-robust graph contrastive learning to reduce the effect of false negative pairs. Finally, we design a cross-view graph alignment to exploit view complementarity. Extensive experiments confirm the superiority of RGCL over existing IMVC methods.

Acknowledgments

This work was supported by the Sichuan Science and Technology Program (Grant nos. 2025ZNSFSC0474, 2024ZDZX0004) , the Mianyang Science and Technology Program (Grant nos. 2023ZYDF003, 2023ZYDF091), the Base Strengthening Program of National Defense Science and Technology (Grant no. 2022-JCJQ-JJ-0292), and the Sichuan Science and Technology Miaozi Program (Grant no. MZGC20240057).

References

[Cai *et al.*, 2012] Xiao Cai, Hua Wang, Heng Huang, and Chris Ding. Joint stage recognition and anatomical annotation of drosophila gene expression patterns. *Bioinformatics*, 28(12):i16–i24, 2012.

[Chao *et al.*, 2024] Guoqing Chao, Yi Jiang, and Dianhui Chu. Incomplete contrastive multi-view clustering with high-confidence guiding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 11221–11229, 2024.

[Geng *et al.*, 2022] Chuanxing Geng, Aiyang Han, and Songcan Chen. View-labels are indispensable: A multifacet complementarity study of multi-view clustering. *arXiv preprint arXiv:2205.02507*, 2022.

[Geusebroek *et al.*, 2005] Jan-Mark Geusebroek, Gertjan J Burghouts, and Arnold WM Smeulders. The amsterdam library of object images. *International Journal of Computer Vision*, 61:103–112, 2005.

[Huang *et al.*, 2024] Binqiang Huang, Zhijie Huang, Shoujie Lan, Qinghai Zheng, and Yuanlong Yu. Incomplete multi-view clustering via inference and evaluation. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 8180–8184, 2024.

[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 *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11600–11609, 2023.

[LeCun *et al.*, 1989] Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.

[Li *et al.*, 2021] Yunfan Li, Peng Hu, Zitao Liu, Dezhong Peng, Joey Tianyi Zhou, and Xi Peng. Contrastive clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 8547–8555, 2021.

[Li *et al.*, 2023a] Haobin Li, Yunfan Li, Mouxing Yang, Peng Hu, Dezhong Peng, and Xi Peng. Incomplete multi-view clustering via prototype-based imputation. *arXiv preprint arXiv:2301.11045*, 2023.

[Li *et al.*, 2023b] Xingfeng Li, Zhenwen Ren, Quansen Sun, and Zhi Xu. Auto-weighted tensor schatten p-norm for robust multi-view graph clustering. *Pattern Recognition*, 134:109083, 2023.

[Li *et al.*, 2023c] 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.*, 2025] Xingfeng Li, Yuangang Pan, Yuan Sun, Quansen Sun, Yinghui Sun, Ivor W. Tsang, and Zhenwen Ren. Incomplete multi-view clustering with paired and balanced dynamic anchor learning. *IEEE Transactions on Multimedia*, 27:1486–1497, 2025.

[Liang *et al.*, 2023] Ke Liang, Yue Liu, Sihang Zhou, Wenxuan Tu, Yi Wen, Xihong Yang, Xiangjun Dong, and Xinwang Liu. Knowledge graph contrastive learning based on relation-symmetrical structure. *IEEE Transactions on Knowledge and Data Engineering*, 36(1):226–238, 2023.

[Liang *et al.*, 2024] Ke Liang, Lingyuan Meng, Meng Liu, Yue Liu, Wenxuan Tu, Siwei Wang, Sihang Zhou, Xinwang Liu, Fuchun Sun, and Kunlun He. A survey of knowledge graph reasoning on graph types: Static, dynamic, and multi-modal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12):9456–9478, 2024.

[Liang *et al.*, 2025] Ke Liang, Lingyuan Meng, Hao Li, Meng Liu, Siwei Wang, Sihang Zhou, Xinwang Liu, and Kunlun He. Mgksite: Multi-modal knowledge-driven site selection via intra and inter-modal graph fusion. *IEEE Transactions on Multimedia*, 27:1722–1735, 2025.

[Lin *et al.*, 2021] Yijie Lin, Yuanbiao Gou, Zitao Liu, Boyun Li, Jiancheng Lv, and Xi Peng. Completer: Incomplete multi-view clustering via contrastive prediction. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11169–11178, 2021.

[Lin *et al.*, 2022] Yijie Lin, Yuanbiao Gou, Xiaotian Liu, Jinfeng Bai, Jiancheng Lv, and Xi Peng. Dual contrastive prediction for incomplete multi-view representation learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4):4447–4461, 2022.

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

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

[Romera-Paredes and Torr, 2015] Bernardino Romera-Paredes and Philip Torr. An embarrassingly simple approach to zero-shot learning. In *International Conference on Machine Learning*, pages 2152–2161, 2015.

[Sun *et al.*, 2023] Yuan Sun, Xu Wang, Dezhong Peng, Zhenwen Ren, and Xiaobo Shen. Hierarchical hashing learning for image set classification. *IEEE Transactions on Image Processing*, 32:1732–1744, 2023.

[Sun *et al.*, 2024] Yuan Sun, Yang Qin, Yongxiang Li, Dezhong Peng, Xi Peng, and Peng Hu. Robust multi-view clustering with noisy correspondence. *IEEE Transactions on Knowledge and Data Engineering*, 36(12):9150–9162, 2024.

[Tang and Liu, 2022] Huayi Tang and Yong Liu. Deep safe incomplete multi-view clustering: Theorem and algorithm. In *International Conference on Machine Learning*, pages 21090–21110. PMLR, 2022.

[Teng *et al.*, 2024] Ge Teng, Ting Mao, Chen Shen, Xiang Tian, Xuesong Liu, Yaowu Chen, and Jieping Ye. Urrl-imvc: Unified and robust representation learning for incomplete multi-view clustering. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 2888–2899, 2024.

[Vassilvitskii and K-means+, 2006] S Vassilvitskii and Arthur D K-means+. The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms.*, pages 1027–1035, 2006.

[Wang *et al.*, 2022] Yiming Wang, Dongxia Chang, Zhiqiang Fu, Jie Wen, and Yao Zhao. Graph contrastive partial multi-view clustering. *IEEE Transactions on Multimedia*, 25:6551–6562, 2022.

[Wang *et al.*, 2023] Jiatai Wang, Zhiwei Xu, Xuewen Yang, Dongjin Guo, and Limin Liu. Self-supervised image clustering from multiple incomplete views via constrastive complementary generation. *IET Computer Vision*, 17(2):189–202, 2023.

[Wang *et al.*, 2024a] Haiyue Wang, Quan Wang, Qiguang Miao, and Xiaoke Ma. Joint learning of data recovering and graph contrastive denoising for incomplete multi-view clustering. *Information Fusion*, 104:102155, 2024.

[Wang *et al.*, 2024b] Haiyue Wang, Wensheng Zhang, and Xiaoke Ma. Contrastive and adversarial regularized multi-level representation learning for incomplete multi-view clustering. *Neural Networks*, 172:106102, 2024.

[Wang *et al.*, 2024c] Yiming Wang, Dongxia Chang, Zhiqiang Fu, Jie Wen, and Yao Zhao. Partially view-aligned representation learning via cross-view graph contrastive network. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(8):7272–7283, 2024.

[Wen *et al.*, 2020] Jie Wen, Zheng Zhang, Yong Xu, Bob Zhang, Lunke Fei, and Guo-Sen Xie. Cdimc-net: Cognitive deep incomplete multi-view clustering network. In *International Joint Conference on Artificial Intelligence*, pages 3220–3236, 2020.

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

[Xu *et al.*, 2022] Jie Xu, Chao Li, Yazhou Ren, Liang Peng, Yujie Mo, Xiaoshuang Shi, and Xiaofeng Zhu. Deep incomplete multi-view clustering via mining cluster complementarity. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 8761–8769, 2022.

[Xu *et al.*, 2023] Jie Xu, Chao Li, Liang Peng, Yazhou Ren, Xiaoshuang Shi, Heng Tao Shen, and Xiaofeng Zhu. Adaptive feature projection with distribution alignment for deep incomplete multi-view clustering. *IEEE Transactions on Image Processing*, 32:1354–1366, 2023.

[Xu *et al.*, 2025] Shilin Xu, Yuan Sun, Xingfeng Li, Siyuan Duan, Zhenwen Ren, Zheng Liu, and Dezhong Peng. Noisy label calibration for multi-view classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 21797–21805, 2025.

[Xue *et al.*, 2021] Zhe Xue, Junping Du, Changwei Zheng, Jie Song, Wenqi Ren, and Meiyu Liang. Clustering-induced adaptive structure enhancing network for incomplete multi-view data. In *International Joint Conference on Artificial Intelligence*, pages 3235–3241, 2021.

[Yang and Newsam, 2010] Yi Yang and Shawn Newsam. Bag-of-visual-words and spatial extensions for land-use classification. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 270–279, 2010.

[Yang *et al.*, 2022] Mouxing Yang, Yunfan Li, Peng Hu, Jinfeng Bai, Jian Cheng Lv, and Xi Peng. Robust multi-view clustering with incomplete information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1):1055–1069, 2022.

[Yang *et al.*, 2023] Xihong Yang, Jin Jiaqi, Siwei Wang, Ke Liang, Yue Liu, Yi Wen, Suyuan Liu, Sihang Zhou, Xinwang Liu, and En Zhu. Dealmvc: Dual contrastive calibration for multi-view clustering. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 337–346, 2023.

[Yuan *et al.*, 2024] Honglin Yuan, Shiyun Lai, Xingfeng Li, Jian Dai, Yuan Sun, and Zhenwen Ren. Robust prototype completion for incomplete multi-view clustering. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pages 10402–10411, 2024.

[Yuan *et al.*, 2025] Honglin Yuan, Yuan Sun, Fei Zhou, Jing Wen, Shihua Yuan, Xiaojian You, and Zhenwen Ren. Prototype matching learning for incomplete multi-view clustering. *IEEE Transactions on Image Processing*, 34:828–841, 2025.

[Zhang *et al.*, 2020] Changqing Zhang, Yajie Cui, Zongbo Han, Joey Tianyi Zhou, Huazhu Fu, and Qinghua Hu. Deep partial multi-view learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(5):2402–2415, 2020.

[Zheng *et al.*, 2021] Mingkai Zheng, Shan You, Fei Wang, Chen Qian, Changshui Zhang, Xiaogang Wang, and Chang Xu. Ressl: Relational self-supervised learning with weak augmentation. *Advances in Neural Information Processing Systems*, 34:2543–2555, 2021.

[Zhu *et al.*, 2025] Zhengzhong Zhu, Chujun Pu, Xuejie Zhang, Jin Wang, and Xiaobing Zhou. Dual-dimensional contrastive learning for incomplete multi-view clustering. *Neurocomputing*, 615:128892, 2025.