Efficient Multi-view Clustering via Reinforcement Contrastive Learning

Qianqian Wang[1] , **Haiming Xu**[1*] , **Zihao Zhang**[1] , **Zhiqiang Tao**[2] and **Quanxue Gao**[1]
[1]School of Telecommunications Engineering, Xidian University, Xi'an, China
[2]Department of Computer Science, Rochester Institute of Technology, Rochester, NY, USA
qqwang@xidian.edu.cn, 24011211044@stu.xidian.edu.cn, zihaoz2021@foxmail.com,
zqtaomail@gmail.com, qxgao@xidian.edu.cn

Abstract

Contrastive multi-view clustering has demonstrated remarkable potential in complex data analysis, yet existing approaches face two critical challenges: difficulty in constructing high-quality positive and negative pairs and high computational overhead due to static optimization strategies. To address these challenges, we propose an innovative **E**fficient **M**ulti-**V**iew **C**lustering framework with **R**einforcement **C**ontrastive **L**earning (EMVC-RCL). Our key innovation is developing a reinforcement contrastive learning paradigm for dynamic clustering optimization. First, we leverage multi-view contrastive learning to obtain latent features, which are then sent to the reinforcement learning module to refine low-quality features. Specifically, it selects high-confident features to guide the positive/negative pair construction of contrastive learning. For the low-confident features, it utilizes the prior balanced distribution to adjust their assignment. Extensive experimental results showcase the effectiveness and superiority of our proposed method on multiple benchmark datasets.

1 Introduction

Multi-view data has become increasingly ubiquitous across diverse real-world applications [Zhan *et al.*, 2018b; Zhang *et al.*, 2016], which describes the same object from multiple perspectives, such as different sensor types [Wang *et al.*, 2015b] and various features [Liu, 2021; Cao *et al.*, 2015b]. Multi-view learning, which can exploit complementary and consistent information across different views, achieves more comprehensive insights than single-view approaches [Li *et al.*, 2019a; Huang *et al.*, 2021; Peng *et al.*, 2022]. As a fundamental multi-view learning paradigm, multi-view clustering (MVC) has emerged as a crucial technique for discovering intrinsic patterns and structures of multi-view data under unsupervised settings [Zhang *et al.*, 2017] and hence receives considerable attention in machine learning and computer vision [Wen *et al.*, 2020; Wang *et al.*, 2018; Zhan *et al.*, 2018a].

Traditional MVC methods primarily obtain consistent latent representation from multi-view data for clustering with linear transformation, which can be roughly categorized into kernel-based methods, subspace-based methods, and nonnegative matrix factorization (NMF) based methods. For example, Gönen and Margolin [2014] proposed a localized data fusion approach for kernel k-means clustering. Liu et al. [2016] introduced matrix-induced regularization to improve the robustness of multiple kernel k-means. Subsequently, Cao et al. [2015a] developed diversity-induced multi-view subspace clustering. Luo et al. [2018] proposed consistent and specific multi-view subspace clustering. However, these traditional approaches rely heavily on predefined kernels or hand-crafted features, which often fail to capture complex semantic relationships and suffer from limited scalability when handling high-dimensional data.

To mitigate these limitations, deep MVC was developed to exploit complex and nonlinear features embedded within multi-view data [Sun *et al.*, 2024]. Typical approaches include autoencoder-based methods, deep CCA, and generative adversarial networks (GAN) based methods [Dong *et al.*, 2020; Andrew *et al.*, 2013; Li *et al.*, 2019b]. Compared to traditional methods, these approaches demonstrate promising clustering performance. However, as unsupervised learning techniques, they lack label guidance and rely solely on feature differences for clustering. This limitation hinders their ability to learn sufficient discriminative information, resulting in suboptimal performance. Some works introduced contrastive learning to enhance discriminative feature by incorporating pseudo labels as guidance [Trosten *et al.*, 2021; Pan and Kang, 2021]. Nevertheless, contrastive MVC methods encounter a situation where the representation quality and clustering accuracy are mutually dependent. Coupled with static and inflexible optimization strategies, these methods incur extensive computational complexity for convergence and limited improvement of clustering performance.

To address the aforementioned challenges, we develop a novel efficient multi-view clustering framework with reinforcement contrastive learning. This framework incorporates a reinforcement learning module to dynamically optimize the clustering process. Specifically, this module select high-quality features for contrastive learning and adaptively refine low-quality features. Then, we propose a memory-enhanced mechanism that leverages historical information,

[*] Corresponding Author

which utilizes temporal feature consistency to efficiently refine ambiguous assignments via momentum-based update. To dynamically optimize the memory-enhanced mechanism, we formulate the parameter adaptation as a reinforcement learning problem and solve it with Q-learning. This approach significantly enhances clustering performance while maintaining computational efficiency. Our main contributions are summarized as follows:

- We propose a novel reinforcement contrastive learning framework for deep MVC, which adaptively selects high-confident features to enhance the discriminative feature learned by contrastive learning.
- We design a memory-enhanced mechanism to store high-quality features and refine low-quality feature via momentum-based update, thereby stabilizing the clustering and promoting faster convergence.
- We conduct extensive experiments and comparisons, demonstrating superior clustering performance with only 30% of the traditional computational cost across various benchmark datasets.

2 Related Work

2.1 Deep Multi-view Clustering

Multi-view clustering (MVC) leverages the consistency and complementarity across views, using similarity measures to partition data into clusters and uncover underlying relationships. Due to the inherent challenges of unsupervised learning, many MVC methods design pretext tasks to extract potential low-dimensional features and shared information across views. For instance, autoencoder-based approaches [Zhang *et al.*, 2019; Fan *et al.*, 2020] enhance feature reliability by enforcing consistency between the original inputs and reconstructed outputs. However, such reconstruction constraints typically capture information within individual views, lacking the ability to explore cross-view correlations. To address this, numerous algorithms employ subspace learning to project multi-view data into a unified shared subspace, enabling joint representation learning [Zhang *et al.*, 2018; Wang *et al.*, 2023b; Wang *et al.*, 2023a; Zhang *et al.*, 2023]. Additionally, some methods integrate CCA constraints [Wang *et al.*, 2015a; Gao *et al.*, 2020] or leverage generative adversarial networks [Li *et al.*, 2019b; Xu *et al.*, 2019] to align feature distributions across different views. Recently, the emergence of contrastive learning has revitalized MVC [Xu *et al.*, 2022; Chen *et al.*, 2023; Wang *et al.*, 2025]. By pulling together positive samples and pushing apart negative ones, contrastive objectives implicitly enhance intra-cluster compactness and inter-cluster separability. For example, Cui *et al.* [2024] proposed a dual contrast-driven clustering framework that improves feature discrimination through contrastive learning. Despite these advancements, existing methods still face key limitations, such as static optimization strategies, suboptimal view-weight balancing, and high computational overhead, all of which hinder their scalability and practical deployment.

2.2 Reinforcement Learning for Clustering

Recently, an emerging line of work has started to explore the potential of reinforcement learning (RL) in multi-view clustering (MVC) [Liu *et al.*, 2023]. Dai *et al.* [2025] pioneered a graph-based framework that employs reinforcement learning to infer the number of clusters, using an inter-cluster connectivity reward mechanism to guide the process. Building on this idea, Gu *et al.* [2023] introduced MVCIR-net, which integrates contrastive learning with reinforcement strategies to enhance the recognition of clustering structures. In a different direction, Yang *et al.* [2022] proposed a self-supervised representation learning framework tailored for multi-view reinforcement learning, aiming to improve sample efficiency and representation quality. While these approaches demonstrate the potential of combining RL with MVC, they tend to focus on isolated objectives—such as cluster number estimation or representation learning—without offering a unified solution. Specifically, they fall short in addressing critical challenges such as class imbalance, adaptive optimization, and computational efficiency in a holistic manner. This underscores the need for a more comprehensive framework that can leverage reinforcement learning not only for structural inference or feature learning but also for enhancing the overall robustness and scalability of MVC methods.

3 Methodology

3.1 Notation & Network Architecture

Given a multi-view dataset $\mathcal{X} = \{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \ldots, \mathbf{X}^{(V)}\}$ with V views and n samples, where $\mathbf{X}^{(v)} \in \mathbb{R}^{n \times d_v}$ represents the feature matrix of view v, our goal is to partition these samples into K clusters. We denote the latent representations as $\mathbf{Z}^{(v)} \in \mathbb{R}^{n \times d_h}$, enhanced features as $\mathbf{H}^{(v)} \in \mathbb{R}^{n \times d_h}$, and clustering probability distributions as $\mathbf{P}^{(v)} \in \mathbb{R}^{n \times K}$, with its regularized version as $\hat{\mathbf{P}}^{(v)}$.

Our EMVC-RCL framework consists of three key modules: View-specific Encoder-Decoder Module with encoders E_v and decoders D_v that preserve view-specific information; Local-Global Contrastive Module that ensures both sample-level alignment and cluster-level consistency across views; and Dynamic Clustering Refinement Module that employs reinforcement learning with memory bank $\mathbf{M}^{(t,v)} \in \mathbb{R}^{n \times d_h}$ at time step t to stabilize uncertain assignments. These modules work together to achieve consistent and balanced clustering while preserving the unique characteristics of each view.

3.2 View-specific Encoder/Decoder

To learn discriminative representations from multi-view data, we employ view-specific encoders and decoders. For each view v, the encoder E_v projects the input features into a latent space:

$$\mathbf{Z}^{(v)} = E_v(\mathbf{X}^{(v)}) \in \mathbb{R}^{n \times d_h} \tag{1}$$

where E_v is implemented as a Multi-Layer Perception (MLP) that maps features from dimension d_v to d_h. Then, we employ the projected features from other views to refine the v-th view feature as follows:

$$\mathbf{H}^{(v)} = \mathbf{Z}^{(v)} + \alpha_t \cdot \frac{1}{V-1} \sum_{u \neq v} \mathbf{Z}^{(u)} \in \mathbb{R}^{n \times d_h} \tag{2}$$

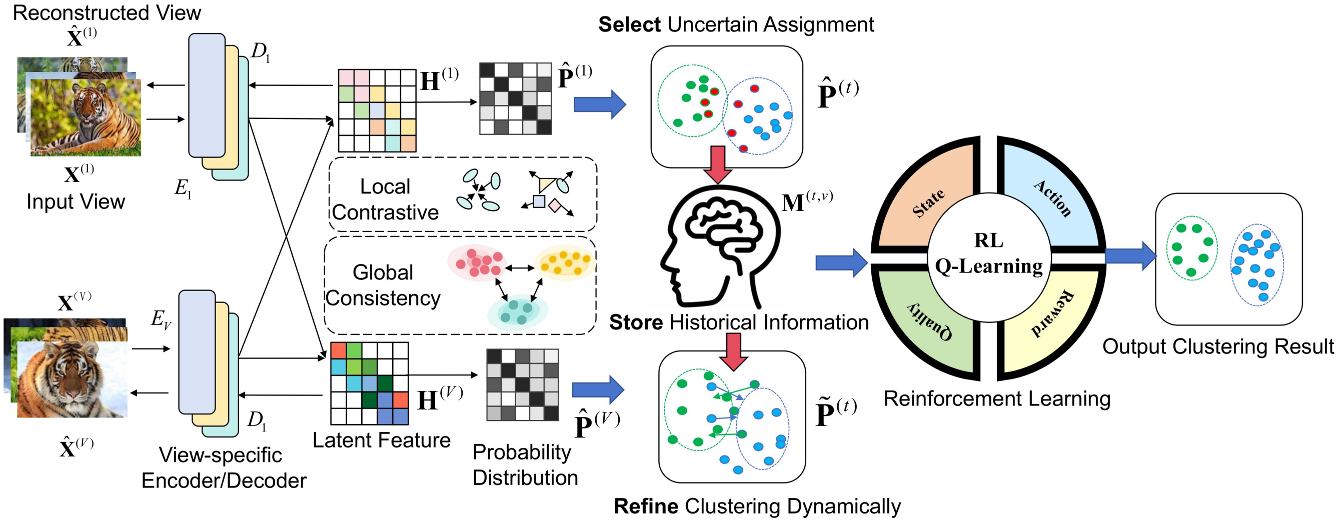


Figure 1: The architecture of our EMVC-RCL framework. The data flow starts from multi-view inputs, which are processed by view-specific encoders and decoders to extract latent features and reconstruct each view. These latent features are used to generate clustering probability distributions, which are further optimized by local contrastive and global consistency objectives. Uncertain assignments are identified and refined using a memory module that stores historical information. Reinforcement learning (Q-Learning) dynamically adjusts clustering strategies by evaluating clustering quality and assignment confidence, ultimately producing clustering results.

where $\alpha_t \in [0, 1]$ controls the intensity of multi-view information integration. To preserve essential view information, we reconstruct the input data through view-specific decoders:

$$\hat{\mathbf{X}}^{(v)} = D_v(\mathbf{H}^{(v)}) \in \mathbb{R}^{n \times d_v} \tag{3}$$

where D_v is also implemented as an MLP. We employ the reconstruction loss to guarantee reconstruction quality:

$$\mathcal{L}_{rec} = \frac{1}{V} \sum_{v=1}^{V} \|\mathbf{X}^{(v)} - \hat{\mathbf{X}}^{(v)}\|_F^2 \tag{4}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. While this reconstruction process helps preserve view-specific information, it primarily focuses on individual view characteristics and may not fully capture the semantic relationships across views.

3.3 Local-Global Contrastive Module

To learn clustering-friendly representations that capture cross-view semantic relationships, we propose a local-global contrastive learning strategy that operates on the enhanced features $\mathbf{H}^{(v)}$ at both sample and cluster levels. The learned features $\mathbf{H}^{(v)}$ are first transformed into clustering probability distributions through a shared network:

$$\mathbf{P}^{(v)} = \varphi(\mathbf{H}^{(v)}) \in \mathbb{R}^{n \times K} \tag{5}$$

where φ consists of two fully-connected layers with softmax activation that map features to probability distributions, and $\mathbf{P}^{(v)}$ represents the probability matrix for view v.

To ensure cross-view consistency at the sample level, we first design a local contrastive loss that operates on individual samples:

$$\mathcal{L}_{local} = \sum_{v=1}^{V} \sum_{u=v+1}^{V} \sum_{i=1}^{n} -\log \frac{\exp(\mathbf{P}_i^{(v)} \cdot \mathbf{P}_i^{(u)} / \tau_1)}{\sum_{j=1}^{n} \exp(\mathbf{P}_i^{(v)} \cdot \mathbf{P}_j^{(u)} / \tau_1)} \tag{6}$$

where $\mathbf{P}_i^{(v)}$ denotes the i-th row of $\mathbf{P}^{(v)}$, representing the clustering probability distribution of sample i in view v, and τ_1 is a temperature parameter controlling the sharpness of the similarity distribution. This local loss treats each sample pair independently, pulling together the same sample's representations across different views while pushing apart different samples.

While the local contrastive loss ensures sample-level alignment, it does not consider the overall cluster distribution, potentially leading to assignment error. To obtain the global cluster distribution structure, we first adjust the cluster distribution to be sharper as follows:

$$\hat{\mathbf{P}}_{ik}^{(v)} = \frac{(\mathbf{P}_{ik}^{(v)})^2 / f_k}{\sum_{k'} (\mathbf{P}_{ik'}^{(v)})^2 / f_{k'}} \tag{7}$$

where $f_k = \sum_{i=1}^{n} \mathbf{P}_{ik}^{(v)}$ represents the size of cluster k, and $\hat{\mathbf{P}}^{(v)}$ is the regularized probability matrix.

Building upon these regularized distributions, we introduce a global consistency loss that aligns the overall cluster structures across views:

$$\mathcal{L}_{global} = \sum_{u \neq v} \mathrm{KL}(\hat{\mathbf{P}}^{(v)} \| \hat{\mathbf{P}}^{(u)}) \tag{8}$$

Unlike the local contrastive loss that focuses on individual samples, the global consistency loss ensures that the entire clustering structure remains consistent across views by minimizing the KL divergence between view-specific cluster distributions. Meanwhile, this loss can avoid sorting most samples assigned into a single cluster. The final local-global contrastive loss combines both local sample-level alignment and global cluster-level consistency:

$$\mathcal{L}_{con} = \mathcal{L}_{local} + \alpha_c \mathcal{L}_{global} \tag{9}$$

where α_c serves as a regularization coefficient controlling the contribution of global cluster structure consistency.

3.4 Dynamic Clustering Refinement Module

We propose a dynamic clustering refinement module with a reinforcement learning framework. Our key insight is that uncertain assignments need special attention and historical information can help stabilize the uncertain assignments and further improve the clustering process.

Selection of Samples with Uncertain Assignment

Definition 1 (Samples with Uncertain Assignment): Given a sample's cluster probability distribution $\hat{\mathbf{P}}_i^{(v)} \in \mathbb{R}^K$, we define the set of uncertain samples as:

$$\mathcal{U}^{(v)} = \{i | \delta_e(\hat{\mathbf{P}}_i^{(v)}) > \eta_e \text{ or } \delta_c(\hat{\mathbf{P}}_i^{(v)}) < \eta_c\} \tag{10}$$

where $\delta_e(\hat{\mathbf{P}}_i^{(v)})$ is the entropy of the assignment $\hat{\mathbf{P}}_i^{(v)}$; $\delta_c(\hat{\mathbf{P}}_i^{(v)})$ is the confidence score; η_e and η_c are thresholds.

To quantify this uncertainty, we evaluate each sample i in view v through two complementary metrics: entropy and confidence score. A sample is considered with an uncertain assignment if it exhibits either high entropy (dispersed distribution) or low confidence (unclear boundary). The entropy of its probability distribution is:

$$\delta_e(\hat{\mathbf{P}}_i^{(v)}) = -\sum_{k=1}^{K} \hat{\mathbf{P}}_{ik}^{(v)} \log \hat{\mathbf{P}}_{ik}^{(v)} \tag{11}$$

where $\hat{\mathbf{P}}_i^{(v)} \in \mathbb{R}^K$ denotes the i-th row of $\hat{\mathbf{P}}^{(v)}$. Additionally, we compute a confidence score as follows:

$$\delta_c(\hat{\mathbf{P}}_i^{(v)}) = \frac{\max_k \hat{\mathbf{P}}_{ik}^{(v)}}{\text{second\_max}_k \hat{\mathbf{P}}_{ik}^{(v)}} \tag{12}$$

A sample with uncertain assignment means its cluster assignments are less reliable. To enhance its clustering quality, we propose to leverage historical information through a memory-based mechanism, as detailed in the next section.

Storage of Historical Information

To stabilize the clustering assignments of samples with uncertain assignments identified in $\mathcal{U}^{(v)}$, we propose a memory-enhanced mechanism for historical information storage, and then, we realize temporal consistency in feature space to refine uncertain assignments through momentum-based updates.

At each time step t, the memory-enhanced mechanism maintains a memory bank $\mathbf{M}^{(t,v)} \in \mathbb{R}^{n \times d_h}$ and updates it through momentum-based dynamics to ensure stable feature patterns:

$$\mathbf{M}^{(t,v)} = \beta \mathbf{M}^{(t-1,v)} + (1-\beta)\mathbf{H}^{(v)} \tag{13}$$

where $\beta \in (0, 1)$ controls the balance between historical memory and current features.

For uncertain samples in $\mathcal{U}^{(v)}$, we establish relationships $\mathcal{S}_{ij}^{(t,v)}$ between i-th current features $\mathbf{H}_i^{(v)}$ and j-th historical memory $\mathbf{M}_j^{(t,v)}$ using cosine similarity:

$$\mathcal{S}_{ij}^{(t,v)} = \frac{\mathbf{H}_i^{(v)} \cdot \mathbf{M}_j^{(t,v)}}{\|\mathbf{H}_i^{(v)}\| \|\mathbf{M}_j^{(t,v)}\|} \tag{14}$$

Then, we compute the contribution weights based on these similarities:

$$\mathbf{w}_{ij} = \frac{\exp(\mathcal{S}_{ij}^{(t,v)}/\tau)}{\sum_{m=1}^{n} \exp(\mathcal{S}_{im}^{(t,v)}/\tau)} \tag{15}$$

where τ is a temperature parameter controlling the sharpness of the weight distribution. The contribution weight $\mathbf{w}_{ij}$ quantifies the influence of historical samples on the current sample's assignment.

Dynamic Clustering Refinement

Definition 2 (Dynamic Sample Assignment): Given the current sample features and historical memory, the adaptive refinement strategy for cluster assignments is defined as:

$$\tilde{\mathbf{P}}_i^{(t)} = \begin{cases} \phi \hat{\mathbf{P}}_i^{(t)} + (1-\phi) \sum_{j=1}^{n} \mathbf{w}_{ij} \hat{\mathbf{P}}_j^{(t-1)}, & \text{if } i \in \mathcal{U}^{(v)} \\ \hat{\mathbf{P}}_i^{(t)}, & \text{otherwise} \end{cases} \tag{16}$$

where $\mathbf{w}_{ij}$ weights the contribution of each memory entry based on its similarity, allowing samples with uncertain assignments to refine their assignments by combining current predictions with reliable historical patterns.

This adaptive refinement strategy stabilizes uncertain assignments but depends on key parameters like τ and β, which need dynamic adjustment during clustering. Thus, we model parameter adaptation as a sequential decision-making problem, explored in the next part.

Parameter Adaptation with Reinforcement Learning

To dynamically update the parameters, we introduce a reinforcement learning (RL) strategy, specifically utilizing Q-learning, which enables an agent to learn optimal decision-making policies by interacting with an environment, where the agent observes **states**, takes **actions**, and receives **rewards** based on those actions.

Definition 3 (State): The state at each time step t is represented as a vector of features:

$$\text{state}^{(t)} = [\Delta_{\text{in}}^{(t)}, \Delta_{\text{out}}^{(t)}, \rho_k^{(t)}] \tag{17}$$

where

$$\begin{cases} \Delta_{\text{in}}^{(t)} = \dfrac{1}{K} \displaystyle\sum_{k=1}^{K} \sum_{i \in \mathcal{C}_k} \|\mathbf{H}_i^{(v)} - \mathbf{c}_k\|_2^2, \\ \Delta_{\text{out}}^{(t)} = \dfrac{1}{K(K-1)} \displaystyle\sum_{k=1}^{K} \sum_{l \neq k} \|\mathbf{c}_k - \mathbf{c}_l\|_2^2, \\ \rho_k^{(t)} = \dfrac{K_{\text{cur}}}{K}. \end{cases}$$

Here, $\Delta_{\text{in}}^{(t)}$ measures the *compactness* of the clusters, where $\mathcal{C}_k$ represents the set of samples in cluster k and $\mathbf{c}_k$ is the corresponding centroid. $\Delta_{\text{out}}^{(t)}$ quantifies the *separation* between clusters, and $\rho_k^{(t)} = \frac{K_{\text{cur}}}{K}$ represents the *progress* of the clustering process, where K_{cur} is the number of completed clusters at time step t, and K is the total number of clusters.

Definition 4 (Action): Based on the observed state, the agent selects parameters for the memory bank as its action, represented as:

$$\text{action}^{(t)} = [\psi, \eta_c, \tau] \tag{18}$$

where ψ controls the balance between current and historical predictions, η_c is the confidence threshold for identifying uncertain samples, and τ is the temperature parameter controlling the sharpness of similarity weights. These parameters govern the memory-based refinement process and influence how uncertain samples utilize historical information.

Definition 5 (Reward): The effectiveness of the selected parameters is evaluated through a reward signal, defined as:

$$\text{reward}^{(t)} = \Delta_{\text{out}}^{(t)} - \sigma_1 \, \Delta_{\text{in}}^{(t)} + \sigma_2 \sum_{i=1}^{n} \mathbf{1}(\delta_c(\hat{\mathbf{P}}_i^{(v)}) > \eta_c) \quad (19)$$

where σ_1 and σ_2 are weighting parameters balancing cluster quality and assignment confidence, and $\mathbf{1}(\cdot)$ is the indicator function that equals 1 when a sample's confidence exceeds the threshold η_c and 0 otherwise, thus counting the total number of confident samples in the summation.

Following the Q-learning framework, we optimize a Q-network that predicts the expected cumulative reward for state-action pairs. Let s, a, and r represent the state, action, and reward, respectively. The Q-network is trained by minimizing the following loss function:

$$\mathcal{L}_{rl} = \mathbb{E}[(r^{(t)} + \gamma \max_a Q(s^{(t+1)}, a) - Q(s^{(t)}, a^{(t)}))^2] \quad (20)$$

where γ is the discount factor. This optimization process enables the agent to learn effective strategies for selecting parameters that refine uncertain samples based on both current and historical clustering information.

3.5 Overall Objective

The overall objective integrates reconstruction, contrastive learning, and reinforcement learning components:

$$\mathcal{L}_{total} = \mathcal{L}_{rec} + \alpha\mathcal{L}_{con} + \lambda\mathcal{L}_{rl} \quad (21)$$

where α and λ are regularization coefficients.

After model convergence, we compute the final clustering assignment using the refined probability distributions $\tilde{\mathbf{P}}^v$ from Eq. (16):

$$y_i = \arg\max_k \left(\frac{1}{V} \sum_{v=1}^{V} \tilde{\mathbf{P}}_{ik}^v \right) \quad (22)$$

This aggregation leverages complementary information from all views to determine the most probable cluster for each sample.

4 Experiment

4.1 Datasets & Metric

We evaluate our method on six real-world multi-view datasets from previous works [Chen *et al.*, 2023; Xu *et al.*, 2023]: COIL-20 with 1,440 grayscale images of 20 objects from different angles, RGB-D containing 1,449 samples with RGB and depth information, BDGP consisting of 2,500 drosophila embryo images with two visual descriptors, Scene-15 including 4,485 scene images with GIST, LBP, and HOG features, MNIST-USPS combining handwritten digit samples from both datasets, and Fashion containing product images with visual, textual, and categorical features. For evaluation, we use three standard metrics: Accuracy (ACC), Normalized Mutual Information (NMI), and Purity (PUR).

4.2 Experiment Setup

To comprehensively evaluate our proposed EMVC-RCL method, we compare it with several state-of-the-art methods: K-means [MacQueen and others, 1967], BSVC, SCAgg [Nie *et al.*, 2014], EE-IMVC [Liu *et al.*, 2020], ASR [Chen *et al.*, 2022], DSIMVC [Tang and Liu, 2022], DSMVC [Tang and Liu, 2022], DCP [Lin *et al.*, 2022], MFLVC [Xu *et al.*, 2022], CVCL [Chen *et al.*, 2023], CMVKL [Wan *et al.*, 2024],ADMC [Zhao *et al.*, 2024]. All experiments were conducted on an NVIDIA GeForce RTX 4090 GPU with CUDA 12.4, implemented using PyTorch (Python 3.10.13).

4.3 Performance Comparison

As shown in Table 1, our EMVC-RCL demonstrates superior performance across all datasets and metrics. On BDGP dataset, our approach achieves 99.64% ACC, 98.65% NMI, and 99.60% PUR, surpassing the second-best method CVCL (99.20% ACC, 97.29% NMI, 99.20% PUR). The performance advantage is particularly evident on challenging datasets like Scene-15, where our method achieves 76.80% ACC, significantly outperforming other competitors. This substantial improvement can be attributed to our reinforcement contrastive learning framework, which effectively selects high-confident features while refining low-quality ones through dynamic clustering optimization. Similar patterns can be observed on MNIST-USPS and Fashion datasets, where our method consistently achieves the best performance (99.98% and 99.58% ACC respectively), demonstrating its robust performance across different data modalities and cluster distributions. These comprehensive results validate the effectiveness of our proposed EMVC-RCL framework in handling diverse multi-view clustering scenarios while maintaining computational efficiency.

4.4 Hyper-parameter Analysis

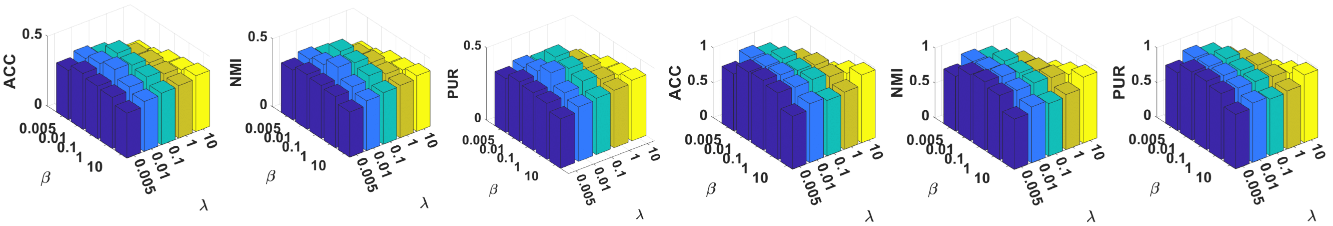
As shown in the Fig. 2, when both λ (geometric relationship consistency) and β (probability distribution consistency) are within the range of [0.01, 1.0], the model exhibits high stability across all three metrics (ACC, NMI, and PUR), indicating that minor perturbations in these hyper-parameters do not significantly affect performance and demonstrating strong generalization ability. When either parameter takes extreme values (such as < 0.005 or > 10), a marked decline in performance is observed, with cluster separation quality being most affected. This pattern is consistent across different datasets, underscoring the critical importance of balancing geometric structure and probability distribution in multi-view clustering. Notably, the performance curves change smoothly with parameter variation, without abrupt fluctuations, further confirming the robustness of our method and its flexibility in practical parameter selection. Moreover, the comparison between the two datasets reflects the universality of our model design and the robustness of its performance.

4.5 Ablation Studies

As shown in Table 2, we conduct comprehensive ablation studies on six benchmark datasets to thoroughly analyze the contribution of each module to the overall performance. The

Methods	COIL-20			RGB-D			BDGP			Scene-15			MNIST-USPS			Fashion		
	ACC	NMI	PUR	ACC	NMI	PUR	ACC	NMI	PUR	ACC	NMI	PUR	ACC	NMI	PUR	ACC	NMI	PUR
K-Means	67.20	78.92	69.72	39.96	36.38	54.11	57.72	47.43	57.72	36.10	36.99	39.40	49.66	44.33	51.70	48.79	48.79	53.66
BSVC	80.21	84.75	80.47	45.34	37.49	51.97	58.05	32.42	54.32	38.05	38.85	42.08	67.98	74.43	72.34	60.32	59.01	63.84
SC$_{Agg}$	73.13	78.46	73.89	44.17	35.10	54.73	38.13	55.71	70.72	39.31	39.31	44.76	89.00	77.12	89.18	98.00	94.80	97.56
EE-IMVC	75.73	83.52	75.76	39.75	27.99	51.62	88.00	71.76	87.76	39.00	33.02	40.27	76.00	68.04	76.48	84.00	79.53	84.45
ASR	80.90	87.60	81.50	23.95	16.94	39.13	97.68	92.63	97.68	42.70	40.17	45.61	97.90	87.72	97.91	96.52	89.94	96.52
CMVKL	32.85	48.00	36.04	21.19	11.20	31.54	77.68	82.99	77.68	31.10	29.31	33.09	50.68	46.04	52.44	57.37	54.50	59.00
DSIMVC	65.55	72.51	66.67	37.13	26.54	40.06	99.04	96.86	99.04	28.27	29.04	29.79	99.34	98.13	99.34	88.21	83.99	88.21
DCP	67.36	78.79	69.86	26.43	22.78	26.43	97.04	92.43	97.04	42.32	40.38	43.85	99.02	97.29	99.02	89.37	88.61	89.37
DSMVC	76.46	84.15	78.19	25.74	14.33	25.17	75.80	61.39	75.80	43.48	41.11	45.92	96.34	94.27	96.34	89.63	86.81	89.63
MFLVC	73.19	81.43	75.07	44.17	25.80	47.55	98.72	96.13	98.72	42.52	40.34	44.53	99.66	99.01	99.66	99.20	98.00	99.20
CVCL	84.65	88.89	85.07	59.52	32.39	48.40	99.20	97.29	99.20	44.59	42.17	47.36	99.70	99.13	99.70	99.31	98.21	99.31
ADMC	77.15	79.43	77.15	75.66	54.27	75.66	96.96	96.12	96.96	43.91	40.57	46.06	91.26	95.50	91.26	89.17	84.92	89.17
Our Method	**95.14**	**98.46**	**95.14**	**80.76**	**71.70**	**80.76**	**99.64**	**98.65**	**99.60**	**44.82**	**47.65**	**48.83**	**99.98**	**99.81**	**99.98**	**99.58**	**96.64**	**99.58**

Table 1: Clustering performance comparison results across all datasets. The optimal result for each metric is presented in boldface, while the second-best result is denoted with an underline.



(a) ACC (b) NMI (c) PUR (d) ACC (e) NMI (f) PUR

Figure 2: Parameter sensitivity analysis on Scene-15 and MNIST-USPS dataset visualized through 3D bar plots. The plots demonstrate how different combinations of λ and β affect the clustering performance metrics.

results clearly demonstrate that all three components $\mathcal{L}_{rec}$, $\mathcal{L}_{con}$, and $\mathcal{L}_{rl}$ are indispensable. With only the reconstruction loss, the model captures basic feature information but achieves limited clustering performance. Introducing the contrastive loss enables the model to better exploit discriminative structures across multiple views, resulting in significant performance gains. Further incorporating the reinforcement learning module leads to optimal results on all metrics, with particularly notable improvements on challenging datasets such as RGB-D and Scene-15. This indicates that the RL module dynamically adjusts clustering strategies and leverages the memory-enhanced mechanism to effectively mitigate the impact of noise and uncertain samples. It is also noteworthy that the contrastive loss plays a pivotal role; its removal causes a substantial drop in performance, underscoring the central importance of multi-view contrastive learning in our framework. Overall, the synergy among these three components not only boosts clustering accuracy but also enhances the model's generalization and robustness, fully validating the rationality and effectiveness of our design.

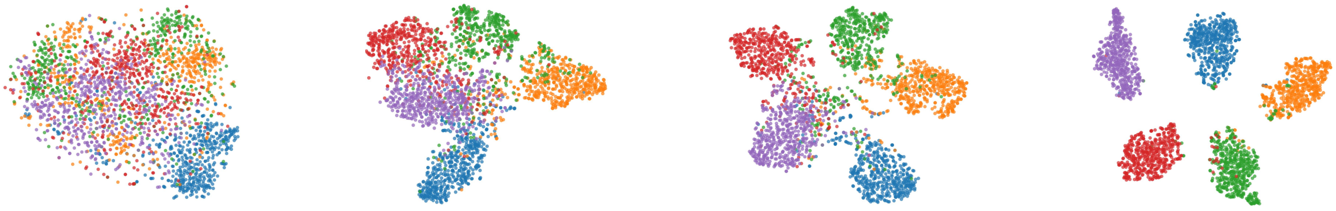
4.6 Visualization Analysis

Fig. 3 illustrates the evolution of feature distributions during the training process using t-SNE visualization. The initial raw features exhibit scattered and mixed clusters with significant overlap. As training progresses, our method gradually refines the feature space: the first stage shows initial cluster formation with reduced inter-cluster overlap, the second stage further enhances cluster separation, and the final stage demonstrates well-defined, compact clusters with clear boundaries. This progressive improvement validates our method's effectiveness in learning discriminative representations through the training process.

Fig. 5 presents a comprehensive comparison between our method and CVCL across six datasets. While maintaining comparable or superior accuracy (represented by bar charts), our method consistently requires significantly less computation time (shown by line plots). For instance, on the MNIST-USPS dataset, our method achieves higher accuracy while reducing the execution time. This substantial improvement in computational efficiency, without compromising clustering performance, demonstrates the practical advantages of our approach.

4.7 Convergence Analysis

Fig. 4 illustrates the convergence behavior of our method. As shown in Fig. 4(a), the training loss exhibits a rapid initial descent and steadily converges after approximately 40 epochs, while the computation time per epoch remains consistently low, demonstrating the computational efficiency of our approach. Notably, the computation time curve shows a slight downward trend, indicating that as the model learns, the reinforcement learning module can more efficiently select samples, reducing the number of uncertain samples that need processing. Fig. 4(b) reveals that both feature quality and class balance metrics improve consistently and stabilize after 60 epochs, validating the effectiveness of our framework in maintaining good feature discrimination while preserving



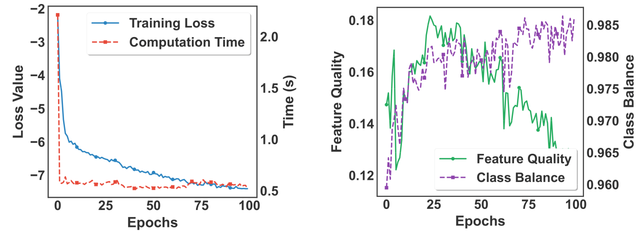
(a) Raw Features (b) 10 Epoch (c) 50 Epoch (d) 100 Epoch

Figure 3: t-SNE visualization of feature evolution on BDGP dataset, demonstrating how features gradually evolve from initial mixed clusters to well-separated final representations through our contrastive learning and feature quality assessment mechanism.

Dataset	$\mathcal{L}_{rec}$	$\mathcal{L}_{con}$	$\mathcal{L}_{rl}$	ACC	NMI	PUR
COIL-20	✓	×	×	73.28	85.15	73.28
	✓	✓	×	83.22	88.04	83.22
	✓	✓	✓	**95.14**	**98.46**	**95.14**
RGB-D	✓	×	×	51.48	23.43	51.48
	✓	✓	×	59.30	32.24	59.30
	✓	✓	✓	**80.76**	**71.70**	**80.76**
BDGP	✓	×	×	74.44	75.30	78.73
	✓	✓	×	97.31	92.65	97.31
	✓	✓	✓	**99.64**	**98.65**	**99.60**
Scene-15	✓	×	×	34.49	38.14	34.49
	✓	✓	×	37.22	40.17	36.58
	✓	✓	✓	**48.82**	**47.65**	**48.83**
MNIST-USPS	✓	×	×	88.01	93.88	88.01
	✓	✓	×	99.12	98.79	99.12
	✓	✓	✓	**99.98**	**99.81**	**99.98**
Fashion	✓	×	×	87.08	84.60	87.08
	✓	✓	×	97.13	94.31	97.13
	✓	✓	✓	**99.58**	**96.64**	**99.58**

Table 2: Ablation study results on six benchmark datasets. The symbols ✓ and × indicate whether each component is enabled: $\mathcal{L}_{rec}$, $\mathcal{L}_{con}$, and $\mathcal{L}_{rl}$.

balanced category distribution. Feature quality and class balance show a slight negative - correlation in fluctuations during optimization, yet both reach high stable levels, evidencing our RL strategy's success in multi - objective optimization.



(a) Loss&Time (b) Quality&Balance

Figure 4: Performance analysis: (a) Convergence of MSE loss during pre-training. (b) Evolution of feature quality and class balance, showing effective learning and balancing.

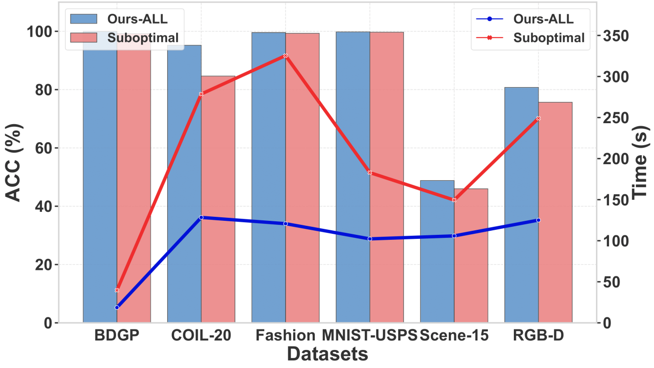


Figure 5: Performance and runtime comparison between our method and the previous best method on each dataset.

5 Conclusions

In this paper, we have presented EMVC-RCL, an innovative efficient multi-view clustering framework with reinforcement contrastive learning, which addresses two critical challenges in contrastive multi-view clustering: the difficulty in constructing high-quality positive and negative pairs, and the high computational overhead caused by static optimization strategies. Our framework introduces a reinforcement contrastive learning paradigm for dynamic clustering optimization, coupled with a memory-enhanced mechanism for efficient feature refinement. Through extensive experiments on six benchmark datasets, we have demonstrated that EMVC-RCL achieves superior clustering performance while requiring only 30% of the traditional training time, validating the effectiveness and efficiency of our approach in complex multi-view data analysis.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant 62176203, the Fundamental Research Funds for the Central Universities (ZYTS25267, QTZX25004), and the Science and Technology Project of Xi'an (Grant 2022JH-JSYF-0009), Open Project of Anhui Provincial Key Laboratory of Multimodal Cognitive Computation, Anhui University (No. MMC202416), Selected Support Project for Scientific and Technological Activities of Returned Overseas Chinese Scholars in Shaanxi Province 2023-02, and the Xidian Innovation Fund (Project NoYJSJ25007).

References

[Andrew *et al.*, 2013] Galen Andrew, Raman Arora, Jeff Bilmes, and Karen Livescu. Deep canonical correlation analysis. In *International conference on machine learning*, pages 1247–1255. PMLR, 2013.

[Cao *et al.*, 2015a] Xiaochun Cao, Changqing Zhang, Huazhu Fu, Si Liu, and Hua Zhang. Diversity-induced multi-view subspace clustering. In *CVPR*, pages 586–594, 2015.

[Cao *et al.*, 2015b] Xiaochun Cao, Changqing Zhang, Chengju Zhou, Huazhu Fu, and Hassan Foroosh. Constrained multi-view video face clustering. *IEEE TIP*, 24(11):4381–4393, 2015.

[Chen *et al.*, 2022] Jie Chen, Shengxiang Yang, Xi Peng, Dezhong Peng, and Zhu Wang. Augmented sparse representation for incomplete multiview clustering. *IEEE TNNLS*, 35(3):4058–4071, 2022.

[Chen *et al.*, 2023] Jie Chen, Hua Mao, Wai Lok Woo, and Xi Peng. Deep multiview clustering by contrasting cluster assignments. in 2023 ieee. In *ICCV*, pages 16706–16715, 2023.

[Cui *et al.*, 2024] Jinrong Cui, Yuting Li, Han Huang, and Jie Wen. Dual contrast-driven deep multi-view clustering. *IEEE TIP*, 2024.

[Dai *et al.*, 2025] Hao Dai, Yang Liu, Peng Su, Hecheng Cai, Shudong Huang, and Jiancheng Lv. Multi-view clustering by inter-cluster connectivity guided reward. In *ICML*, 2025.

[Dong *et al.*, 2020] Shihao Dong, Huiying Xu, Xinzhong Zhu, XiFeng Guo, Xinwang Liu, and Xia Wang. Multi-view deep clustering based on autoencoder. In *JPCS*, volume 1684, page 012059. IOP Publishing, 2020.

[Fan *et al.*, 2020] Shaohua Fan, Xiao Wang, Chuan Shi, Emiao Lu, Ken Lin, and Bai Wang. One2multi graph autoencoder for multi-view graph clustering. In *WWW*, pages 3070–3076, 2020.

[Gao *et al.*, 2020] Quanxue Gao, Huanhuan Lian, Qianqian Wang, and Gan Sun. Cross-modal subspace clustering via deep canonical correlation analysis. In *AAAI*, volume 34, pages 3938–3945, 2020.

[Gönen and Margolin, 2014] Mehmet Gönen and Adam A Margolin. Localized data fusion for kernel k-means clustering with application to cancer biology. *Advances in neural information processing systems*, 27, 2014.

[Gu *et al.*, 2023] Shaokui Gu, Xu Yuan, Liang Zhao, Zhenjiao Liu, Yan Hu, and Zhikui Chen. Mvcir-net: Multi-view clustering information reinforcement network. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 3609–3618, 2023.

[Huang *et al.*, 2021] Shudong Huang, Ivor W Tsang, Zenglin Xu, and Jiancheng Lv. Measuring diversity in graph learning: A unified framework for structured multi-view clustering. *IEEE TKDE*, 34(12):5869–5883, 2021.

[Li *et al.*, 2019a] Ruihuang Li, Changqing Zhang, Huazhu Fu, Xi Peng, Tianyi Zhou, and Qinghua Hu. Reciprocal multi-layer subspace learning for multi-view clustering. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 8172–8180, 2019.

[Li *et al.*, 2019b] Zhaoyang Li, Qianqian Wang, Zhiqiang Tao, Quanxue Gao, Zhaohua Yang, et al. Deep adversarial multi-view clustering network. In *IJCAI*, volume 2, page 4, 2019.

[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 TPAMI*, 45(4):4447–4461, 2022.

[Liu *et al.*, 2016] Xinwang Liu, Yong Dou, Jianping Yin, Lei Wang, and En Zhu. Multiple kernel k-means clustering with matrix-induced regularization. In *AAAI*, volume 30, 2016.

[Liu *et al.*, 2020] Xinwang Liu, Miaomiao Li, Chang Tang, Jingyuan Xia, Jian Xiong, Li Liu, Marius Kloft, and En Zhu. Efficient and effective regularized incomplete multi-view clustering. *IEEE TPAMI*, 43(8):2634–2646, 2020.

[Liu *et al.*, 2023] Yue Liu, Ke Liang, Jun Xia, Xihong Yang, Sihang Zhou, Meng Liu, Xinwang Liu, and Stan Z Li. Reinforcement graph clustering with unknown cluster number. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 3528–3537, 2023.

[Liu, 2021] Xinwang Liu. Incomplete multiple kernel alignment maximization for clustering. *IEEE TPAMI*, 46(3):1412–1424, 2021.

[Luo *et al.*, 2018] Shirui Luo, Changqing Zhang, Wei Zhang, and Xiaochun Cao. Consistent and specific multi-view subspace clustering. In *AAAI*, volume 32, 2018.

[MacQueen and others, 1967] James MacQueen et al. Some methods for classification and analysis of multivariate observations. In *BSMSP*, volume 1, pages 281–297. Oakland, CA, USA, 1967.

[Nie *et al.*, 2014] Feiping Nie, Xiaoqian Wang, and Heng Huang. Clustering and projected clustering with adaptive neighbors. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 977–986, 2014.

[Pan and Kang, 2021] Erlin Pan and Zhao Kang. Multi-view contrastive graph clustering. *Advances in neural information processing systems*, 34:2148–2159, 2021.

[Peng *et al.*, 2022] Liang Peng, Rongyao Hu, Fei Kong, Jiangzhang Gan, Yujie Mo, Xiaoshuang Shi, and Xiaofeng Zhu. Reverse graph learning for graph neural network. *IEEE TNNLS*, 2022.

[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 *ICML*, pages 21090–21110. PMLR, 2022.

[Trosten *et al.*, 2021] Daniel J Trosten, Sigurd Lokse, Robert Jenssen, and Michael Kampffmeyer. Reconsidering representation alignment for multi-view clustering. In *CVPR*, pages 1255–1265, 2021.

[Wan *et al.*, 2024] Xinhang Wan, Bin Xiao, Xinwang Liu, Jiyuan Liu, Weixuan Liang, and En Zhu. Fast continual multi-view clustering with incomplete views. *IEEE TIP*, 33:2995–3008, 2024.

[Wang *et al.*, 2015a] Weiran Wang, Raman Arora, Karen Livescu, and Jeff Bilmes. On deep multi-view representation learning. In *ICML*, pages 1083–1092, 2015.

[Wang *et al.*, 2015b] Yang Wang, Xuemin Lin, Lin Wu, Wenjie Zhang, Qing Zhang, and Xiaodi Huang. Robust subspace clustering for multi-view data by exploiting correlation consensus. *IEEE TIP*, 24(11):3939–3949, 2015.

[Wang *et al.*, 2018] Yang Wang, Lin Wu, Xuemin Lin, and Junbin Gao. Multiview spectral clustering via structured low-rank matrix factorization. *IEEE TNNLS*, 29(10):4833–4843, 2018.

[Wang *et al.*, 2023a] Jing Wang, Songhe Feng, Gengyu Lyu, and Zhibin Gu. Triple-granularity contrastive learning for deep multi-view subspace clustering. In *ACMMM*, pages 2994–3002, 2023.

[Wang *et al.*, 2023b] Shiye Wang, Changsheng Li, Yanming Li, Ye Yuan, and Guoren Wang. Self-supervised information bottleneck for deep multi-view subspace clustering. *IEEE TIP*, 32:1555–1567, 2023.

[Wang *et al.*, 2025] Qianqian Wang, Zihao Zhang, Wei Feng, Zhiqiang Tao, and Quanxue Gao. Contrastive multi-view subspace clustering via tensor transformers autoencoder. In *AAAI*, volume 39, pages 21207–21215, 2025.

[Wen *et al.*, 2020] Jie Wen, Zheng Zhang, Zhao Zhang, Lunke Fei, and Meng Wang. Generalized incomplete multiview clustering with flexible locality structure diffusion. *IEEE TCYB*, 51(1):101–114, 2020.

[Xu *et al.*, 2019] Cai Xu, Ziyu Guan, Wei Zhao, Hongchang Wu, Yunfei Niu, and Beilei Ling. Adversarial incomplete multi-view clustering. In *IJCAI*, volume 7, pages 3933–3939, 2019.

[Xu *et al.*, 2022] Jie Xu, Huayi Tang, Yazhou Ren, Liang Peng, Xiaofeng Zhu, and Lifang He. Multi-level feature learning for contrastive multi-view clustering. In *CVPR*, pages 16051–16060, 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 TIP*, 32:1354–1366, 2023.

[Yang *et al.*, 2022] Huanhuan Yang, Dianxi Shi, Guojun Xie, Yingxuan Peng, Yi Zhang, Yantai Yang, and Shaowu Yang. Self-supervised representations for multi-view reinforcement learning. In *The 38th Conference on Uncertainty in Artificial Intelligence*, 2022.

[Zhan *et al.*, 2018a] Kun Zhan, Feiping Nie, Jing Wang, and Yi Yang. Multiview consensus graph clustering. *IEEE TIP*, 28(3):1261–1270, 2018.

[Zhan *et al.*, 2018b] Kun Zhan, Chaoxi Niu, Changlu Chen, Feiping Nie, Changqing Zhang, and Yi Yang. Graph structure fusion for multiview clustering. *IEEE TKDE*, 31(10):1984–1993, 2018.

[Zhang *et al.*, 2016] Changqing Zhang, Huazhu Fu, Qinghua Hu, Pengfei Zhu, and Xiaochun Cao. Flexible multi-view dimensionality co-reduction. *IEEE TIP*, 26(2):648–659, 2016.

[Zhang *et al.*, 2017] Changqing Zhang, Qinghua Hu, Huazhu Fu, Pengfei Zhu, and Xiaochun Cao. Latent multi-view subspace clustering. In *CVPR*, pages 4279–4287, 2017.

[Zhang *et al.*, 2018] Changqing Zhang, Huazhu Fu, Qinghua Hu, Xiaochun Cao, Yuan Xie, Dacheng Tao, and Dong Xu. Generalized latent multi-view subspace clustering. *IEEE TPAMI*, 42(1):86–99, 2018.

[Zhang *et al.*, 2019] Changqing Zhang, Yeqing Liu, and Huazhu Fu. Ae2-nets: Autoencoder in autoencoder networks. In *CVPR*, pages 2577–2585, 2019.

[Zhang *et al.*, 2023] Zihao Zhang, Qianqian Wang, Zhiqiang Tao, Quanxue Gao, and Wei Feng. Dropping pathways towards deep multi-view graph subspace clustering networks. In *ACMMM*, pages 3259–3267, 2023.

[Zhao *et al.*, 2024] Helin Zhao, Wei Chen, and Peng Zhou. Active deep multi-view clustering. In *IJCAI*, pages 5554–5562, 2024.