Anti-Degradation Lifelong Multi-View Clustering

Xingfeng Li[1], Hao Pan[1], Honglin Yuan[1], Yuan Sun[2]*, Xujian Zhao[1], Jiaqi Lin[3], Zhenwen Ren[1,*]
[1] Southwest University of Science and Technology, Mianyang, China 621010
[2] Sichuan University, Chengdu, China 610044
[3] Agency for Science, Technology and Research, Singapore 138632
lixingfeng@njust.edu.cn, pon915@163.com, hl_yuan0822@163.com, sunyuan_work@163.com,
jasonzhaoxj@swust.edu.cn, linjq56@mail2.sysu.edu.cn, rzw@njust.edu.cn

Abstract

*In real-world scenarios, new views are continuously collected over time, forming a dynamic view stream. To handle such evolving data, a lifelong multi-view clustering framework is needed instead of a static model. However, large discrepancies across views make it challenging to learn new knowledge while preserving previously acquired information. There are few methods use consistency alignment or knowledge distillation to align new knowledge with old ones. However, these strategies cannot fundamentally prevent knowledge degradation, since new knowledge inevitably interferes with the learned representation space. To overcome this limitation, we propose a new **A**nti-degradation **L**ifelong **M**ulti-view **C**lustering (**ALMC**) framework. Specifically, we propose a null-space–projection anti-degradation mechanism that restricts new updates to orthogonal directions of the retained knowledge, ensuring non-interference and preventing degradation; its validity is further supported by experimental and theoretical proof. Extensive experiments on multiple multi-view benchmark datasets demonstrate superior performance in multi-view clustering. The code is available at* https://github.com/lee-xingfeng/ALMC/.

1. Introduction

In data proliferation era, heterogeneous modalities provide complementary perspectives, whereas relying on a single source often fails to capture the inherent complexity and diversity of real-world data [6, 23, 24, 28, 38]. For instance, accurate medical diagnosis typically requires integrating physiological measurements, clinical records, and imaging, since isolating any modality may obscure crucial information and lead to biased decisions. To alleviate such view-specific bias, Multi-View Clustering (MVC) jointly uncovers shared structures across diverse sources by fully exploiting their complementary properties [5, 6, 24, 25, 39, 44, 50, 52, 57, 57]. From the perspective of representation space, existing MVC methods could be divided into four main categories [19]: (1) Multi-view subspace clustering models shared structures in feature subspaces [17, 18, 20, 28, 42, 47, 49]; (2) Multi-view graph clustering leverages semantic relations in graph space [7, 12, 15, 16, 41, 54, 56]; (3) Multiple kernel clustering integrates heterogeneous similarities in the kernel Hilbert space [6, 23, 24, 57]; and (4) Deep multi-view clustering captures consensus patterns in deep representation space [1, 3, 5, 6, 24, 25, 32, 39, 50, 52, 57].

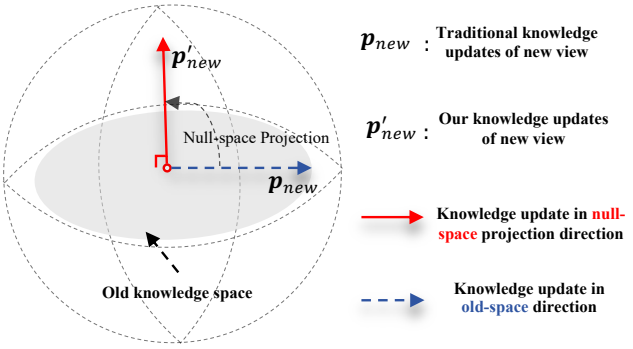


Figure 1. Knowledge anti-degradation mechanism. Traditional methods apply new-view updates along the old knowledge space (blue arrow), causing the new prototypes to slide within this space and overwrite or distort previously learned information. In contrast, ALMC projects each incoming update into a null-space direction (red arrow) that is orthogonal to old knowledge space, preventing interference and preserving all historical knowledge.

Among these paradigms, deep multi-view clustering has gained considerable attention due to its strong ability to learn nonlinear representations. Despite their effectiveness, most existing methods implicitly assume that all views are fully accessible beforehand. However, real-world applications often violate this assumption, since multi-view data

*Corresponding authors.

are typically collected in a continuous and dynamically evolving manner rather than as a fixed dataset. For example, in a smart surveillance system, video, thermal, and audio streams are continuously gathered from distributed sensors; while any single modality may become unreliable under low illumination, occlusion, or background noise, jointly leveraging these evolving views enables more robust and accurate anomaly prediction [29]. To handle continuously streaming multi-view data, several efforts have been made. LSVC [14] introduces a lifelong learning framework based on rule-based updates in a latent knowledge library, but its reliance on preset rules limits its robustness under distribution shifts. DSVC [53] further aligns prototype distributions across views to mitigate incremental drift. Although both methods aim to extract useful information while suppressing redundancy and noise, neither guarantees that newly incorporated knowledge will preserve previously learned representations. Since the knowledge of past views is preserved only in a compressed prototype, incremental updates may further blur or mislead this earlier information. These biased prototypes are then used to guide new views, and the errors accumulate throughout lifelong learning, eventually leading to severely degraded and unreliable clustering results. As illustrated in Fig. 1 and Fig. 6b, LSVC exhibits clear degradation when new views overwrite existing knowledge within the same representation space. Overall, such cumulative **knowledge degradation** poses a fundamental challenge for lifelong streaming multi-view clustering and calls for effective mitigation.

To handle above urgent knowledge degradation issue, we innovatively design an **A**nti-degradation **L**ifelong **M**ulti-view **C**lustering (**ALMC**) framework for streaming view data. Concretely, ALMC prevents knowledge interference by projecting each new-view prototype onto the null space of the subspace spanned by existing knowledge, ensuring orthogonal updates that prevent overwriting past knowledge. A lightweight refinement stabilizes prototype evolution, and a prototype-guided loss aligns representations with both current and stored prototypes. For each new streaming view, its features are projected into the null space to to prevent interference with earlier views. Together, these components form a robust lifelong multi-view clustering framework with stable performance growth across streaming views. Our main contributions are as follows:

- We propose a new anti-degradation lifelong multi-view learning paradigm that explicitly embeds knowledge degradation prevention with theoretical proof into streaming-view modeling. This paradigm provides a principled reference for the streaming multi-view lifelong learning community and offers transferable guidance for designing knowledge-preserving clustering and representation methods.
- Building upon this paradigm, we innovatively design an **A**nti-degradation **L**ifelong **M**ulti-view **C**lustering (**ALMC**) framework in streaming view scenario. To the best of our knowledge, this framework could be the first work to reveal and study the knowledge degradation issue for streaming view clustering.
- We theoretically and experimentally validate the anti-degradation capability of the proposed ALMC, and extensive experiments further show that our ALMC consistently outperforms existing state-of-the-art competitors.

2. Related Work

2.1. Multi-view Clustering

From the perspective of representation space, multi-view clustering [1, 3, 7, 33, 52, 54] can be organized into four paradigms. First in subspace space, multi-view subspace clustering assumes a shared low-dimensional structure and explains each view via sparse/low-rank compositions, yielding robustness to corruption and missing entries while recovering localized substructures and modeling uncertainty in a shared code. Second in graph semantic space, multi-view graph clustering operates on view-wise and consensus affinity graphs, enforcing reliable neighborhoods and manifold agreement while correcting noisy cross-view matches and avoiding fragile spectral dependencies [5, 32, 42, 50]. Third in kernel RKHS space, multiple kernel clustering integrates heterogeneous similarities through kernel weighting or tensorized combinations, with scalability via min–max robust fusion, Nyström approximations, and federated aggregation for privacy-preserving large-scale scenarios [6, 23, 24, 57]. Finally in deep representation space, deep multi-view clustering jointly learns feature extractors and clustering objectives to couple imputation, alignment, and robustness; beyond data-space completion [22], generative modules synthesize missing views [43], cooperation is guided by view-contribution estimation and anchor-shift mitigation [25, 39], robustness is enforced by contrastive training under noisy correspondences [10], and open-world settings leverage correlation-guided novel-class discovery [36]. Prototype-centric strategies further relate samples to learned prototypes to impute absent semantics and stabilize assignments under incompleteness [51]. Despite this progress, most pipelines still assume static views and stable correspondences; real deployments feature streaming, view-asynchronous data and tight resources motivate our deep streaming view clustering framework for dynamic streaming views in real-world environments.

2.2. Lifelong Learning

Lifelong learning is a paradigm that processes data arriving sequentially and updates a compact repository of learned knowledge over time. In lifelong clustering, this repository (*e.g.*, prototypes, anchors, consensus structures) must

preserve established partitions while accommodating new evidence under limited memory and shifting distributions. Prior work spans two settings: *sample-stream* methods that stabilize spectral structures or employ dual memories to retain historical separability [30, 31, 48, 58], and *view-stream* methods that align each newly arrived view to a historical repository and then refine anchors/graphs or distill features [34, 35, 37, 55]. Shallow LSVC offers a lightweight baseline for stream-view fusion [14], while deep AdaptCMVC emphasizes robust consolidation under incremental views [40]; DSVC [53] further instantiates a deep streaming-view pipeline centered on knowledge aggregation, distribution consistency, and knowledge-guided updates.

However, existing formulations [14, 30, 31, 34, 35, 37, 40, 48, 55, 58] pay limited attention to how the knowledge base is constructed and updated. In practice: (1) biased signals accumulate and contaminate prototypes and statistics; (2) cross-view mismatch induces prototype drift or collapse; and (3) tight memory budgets suppress rare or weak-quality views, jointly yielding unstable assignments and a bias toward recent data. We address these issues with a *lifelong anti-degradation learning* paradigm for the knowledge base: each new-view knowledge update is projected onto the null space, where new knowledge in knowledge library is orthogonal to the historical discriminative knowledge, ensuring zero interference with established structure and preserving inter-prototype geometry and decision margins. This paradigm complements spectral, anchor, and contrastive consolidation strategies by enabling reliable knowledge accumulation and robust clustering over long-horizon view streams.

3. Method

3.1. Lifelong Anti-Degradation Mechanism

Observing from Fig. 1, ensuring knowledge preservation is a key challenge in lifelong learning. We tackle this by incorporating mathematical null space projection, establishing a lifelong anti-degradation learning paradigm.

Given an operator A, its *null space* is the set of vectors v that satisfy $Av = 0$, *i.e.*,

$$\text{Null}(\mathbf{A}) = \{\mathbf{v} \in \mathbf{V} : \mathbf{A}\mathbf{v} = \mathbf{0}\}. \tag{1}$$

To facilitate understanding, we will introduce this paradigm into lifelong MVC for illustration. In the knowledge space, when view $v > 1$, let the stored prototype knowledge be denoted as $\mathbf{P}^B$ and the current-view prototype as $\mathbf{P}^n$. To project the current prototype onto the null space of $\mathbf{P}^B$, we first compute the projection matrix of $\mathbf{P}^B$.

Lemma 1. *Let $\mathbf{P}^B$ be an $m \times n$ matrix. Then $\mathbf{P}^B$ and $\mathbf{P}^B(\mathbf{P}^B)^T$ share the same left null space.*

Proof. Define the left null space of a matrix $\mathbf{P}^B$ as

$$\mathcal{N}_L(\mathbf{P}^B) = \left\{\mathbf{x} : \mathbf{x}^T\mathbf{P}^B = \mathbf{0}\right\}. \tag{2}$$

We aim to prove that

$$\mathcal{N}_L(\mathbf{P}^B) = \mathcal{N}_L\left(\mathbf{P}^B(\mathbf{P}^B)^T\right). \tag{3}$$

($\subseteq$) Suppose $\mathbf{x} \in \mathcal{N}_L(\mathbf{P}^B)$, *i.e.*, $\mathbf{x}^T\mathbf{P}^B = \mathbf{0}$. Then,

$$\mathbf{x}^T\mathbf{P}^B(\mathbf{P}^B)^T = (\mathbf{x}^T\mathbf{P}^B)(\mathbf{P}^B)^T = \mathbf{0} \cdot (\mathbf{P}^B)^T = \mathbf{0}. \tag{4}$$

Hence, $\mathbf{x} \in \mathcal{N}_L\left(\mathbf{P}^B(\mathbf{P}^B)^T\right)$, proving $\mathcal{N}_L(\mathbf{P}^B) \subseteq \mathcal{N}_L(\mathbf{P}^B(\mathbf{P}^B)^T)$.

($\supseteq$) Conversely, suppose $\mathbf{x} \in \mathcal{N}_L(\mathbf{P}^B(\mathbf{P}^B)^T)$, *i.e.*,

$$\mathbf{x}^T\mathbf{P}^B(\mathbf{P}^B)^T = \mathbf{0}. \tag{5}$$

Let $\mathbf{y}^T = \mathbf{x}^T\mathbf{P}^B$. Then the above equation becomes$\mathbf{y}^T(\mathbf{P}^B)^T = \mathbf{0}$. This implies that $\mathbf{y}^T = \mathbf{0}$, since the only solution to $\mathbf{y}^T(\mathbf{P}^B)^T = \mathbf{0}$ is $\mathbf{y} = \mathbf{0}$. Hence,

$$\mathbf{x}^T\mathbf{P}^B = \mathbf{y}^T = \mathbf{0}, \tag{6}$$

which shows that $\mathbf{x} \in \mathcal{N}_L(\mathbf{P}^B)$. Therefore, we obtain

$$\mathcal{N}_L(\mathbf{P}^B(\mathbf{P}^B)^T) \subseteq \mathcal{N}_L(\mathbf{P}^B). \tag{7}$$

Combining both inclusions yields $\mathcal{N}_L(\mathbf{P}^B) = \mathcal{N}_L\left(\mathbf{P}^B(\mathbf{P}^B)^T\right)$. □

For computational simplicity, we derive the projection matrix from $\mathbf{P}^B(\mathbf{P}^B)^T \in \mathbb{R}^{d\times d}$, whose null space is equivalent to that of $\mathbf{P}^B$ (the proof is in Lemma 1). We perform Singular Value Decomposition (SVD) on $\mathbf{P}^B(\mathbf{P}^B)^T$ to obtain its eigen-decomposition:

$$\left\{\mathbf{U}, \mathbf{\Lambda}, \mathbf{U}^T\right\} = \text{SVD}\left(\mathbf{P}^B(\mathbf{P}^B)^T\right), \tag{8}$$

where each column of $\mathbf{U}$ corresponds to an eigenvector of $\mathbf{P}^B(\mathbf{P}^B)^T$. By removing the columns associated with non-zero eigenvalues, we obtain the submatrix $\hat{\mathbf{U}}$, and define the projection matrix as:

$$\mathbf{P} = \hat{\mathbf{U}}\hat{\mathbf{U}}^T. \tag{9}$$

The projected new prototype is then given by

$$\mathbf{q}^\perp = \mathbf{P}\mathbf{P}^n \in \mathcal{S}^\perp, \tag{10}$$

where $\mathcal{S}^\perp$ denotes the orthogonal complement of the old prototype subspace $\mathcal{S}$, ensuring that the projected prototype $\mathbf{q}^\perp$ contains information not represented in the existing knowledge space.

Lemma 2. *Let $\mathbf{q}^\perp = \hat{\mathbf{U}}\hat{\mathbf{U}}^T\mathbf{P}^n$. For any $\mathbf{p} \in \mathcal{S}$ (where $\mathcal{S}$ denotes the linear subspace spanned by the old prototypes in the knowledge space), it holds that $\mathbf{p}^T\mathbf{q}^\perp = \mathbf{0}$.*

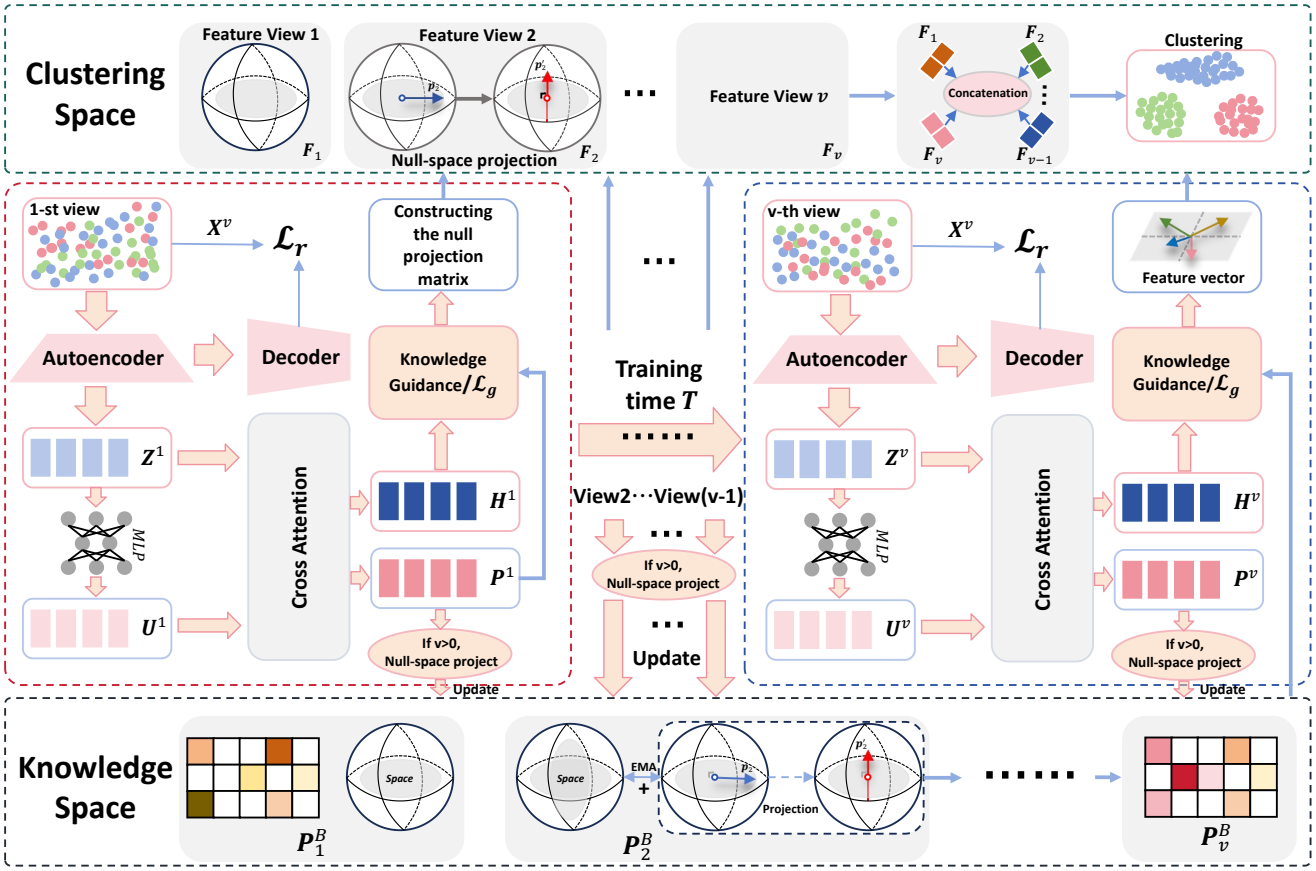


Figure 2. The framework of ALMC. We extract latent features through an autoencoder, then reconstruct them using a decoder, and constrain clustering results with a reconstruction loss function and view features to prevent redundancy from affecting clustering results. An attention mechanism is used to enhance salient features and prototypes, which are then projected onto knowledge space and clustering space via null projection. When $v = 1$, the current prototype is used as a guide for features; when $v > 1$, the knowledge base prototype is used.

Proof. Since $\mathbf{P}^B \in \mathcal{S}$, there exists a coefficient vector $\mathbf{a}$ such that $\mathbf{p} = (\mathbf{P}^B)^T\mathbf{a}$. Then,

$$\mathbf{p}^T\mathbf{q}^\perp = \mathbf{p}^T\hat{\mathbf{U}}\hat{\mathbf{U}}^T\mathbf{P}^n = \left(\hat{\mathbf{U}}^T\mathbf{p}\right)^T\left(\hat{\mathbf{U}}^T\mathbf{P}^n\right). \quad (11)$$

Since the columns of $\hat{\mathbf{U}}$ lie in the orthogonal complement of $\mathcal{S}$, we have $\hat{\mathbf{U}}^T\mathbf{p} = \mathbf{0}$, and thus $\mathbf{p}^T\mathbf{q}^\perp = \mathbf{0}$. □

For any $\mathbf{p} \in \mathcal{S}$, it holds that $\mathbf{p}^T\mathbf{q}^\perp = \mathbf{0}$, as proved in Lemma 2, so $\mathbf{q}^\perp$ is orthogonal to the old prototypes and provides new knowledge in a novel direction. During the fusion stage between old and new prototypes, we adopt an EMA-based weighted fusion strategy, formulated as:

$$\mathbf{P}^{B'} = (1-\alpha)\mathbf{P}^B + \alpha\mathbf{q}^\perp. \quad (12)$$

Projecting $\mathbf{P}^{B'}$ back onto $\mathcal{S}$ using the projection operator $P_{\mathcal{S}}(\cdot)$ yields:

$$\begin{aligned} P_{\mathcal{S}}(\mathbf{P}^{B'}) &= P_{\mathcal{S}}\left((1-\alpha)\mathbf{P}^B + \alpha\mathbf{q}^\perp\right) \\ &= (1-\alpha)P_{\mathcal{S}}(\mathbf{P}^B) + \alpha P_{\mathcal{S}}(\mathbf{q}^\perp) \\ &= (1-\alpha)\mathbf{P}^B. \end{aligned} \quad (13)$$

Over multiple updates, the knowledge space accumulates new information along the null-space direction, formalized as:

$$\mathbf{P}_\mathbf{t}^\mathbf{B} = (1-\alpha)^t\mathbf{P}^B + \sum_{k=0}^{t-1}(1-\alpha)^k\mathbf{q}_{t-k}^\perp. \quad (14)$$

Thus, the fused prototypes continue to lie within the original subspace spanned by $\mathbf{P}^B$, without altering its direction. The EMA update merely scales the magnitude of $\mathbf{P}^B$ by $(1-\alpha)$, a fully controllable factor. Moreover, since $\mathbf{q}^\perp$ is orthogonal to $\mathcal{S}$, its projection component on the old subspace is zero, ensuring that the old prototype representation remains unaffected.

3.2. Lifelong Anti-Degeneration Clustering

3.2.1. Overview

We assume that the multi-view data satisfy $\{\mathbf{X}^i\}_{i=1}^V \in \mathbb{R}^{N\times D^v}$, where the dataset contains V views, each with a feature dimension D^v, and consists of N samples in total.

As illustrated in the Fig. 2, each view is first encoded by an autoencoder to extract the latent feature representations $\mathbf{Z}^v \in \mathbb{R}^{N\times d}$, and we unify their dimensions as d for subsequent processing. To capture both shared and view-specific information, we map each latent feature $\mathbf{Z}^v$ through an MLP to obtain the corresponding prototype $\mathbf{U}^v$. However, the MLP alone cannot effectively extract shared knowledge from complex feature representations. Thus, we introduce a cross-attention mechanism that derives more discriminative prototype knowledge and feature representations from $\mathbf{Z}^v$ and $\mathbf{U}^v$. We formulate this process as:

$$\mathbf{A}^v = \text{Softmax}\left(\frac{(\mathbf{W}_Z^v\mathbf{Z}^v)^T\mathbf{W}_U^v\mathbf{U}^v}{\sqrt{d}}\right), \tag{15}$$

$$\mathbf{P}^v = \mathbf{U}^v + \mathbf{A}^v\mathbf{W}_H'^v\mathbf{Z}^v, \tag{16}$$

$$\mathbf{H}^v = \mathbf{Z}^v + \mathbf{A}^v\mathbf{W}_P'^v\mathbf{U}^v, \tag{17}$$

where $\mathbf{W}_Z^v$ and $\mathbf{W}_U^v$ denote the linear mapping layers for the latent features and prototypes of the v-th view, respectively. $\mathbf{W}_H'^v$ and $\mathbf{W}_P'^v$ represent additional linear transformation layers that refine the prototype and feature representations, respectively.

In the context of streaming view learning. We build a knowledge space to store the prototype knowledge accumulated from the view stream, ensuring that all prototypes are effectively utilized. However, this inevitably introduces a new challenge, *e.g.*, **knowledge degradation**.

To address this issue, the Lifelong Anti-Degradation Mechanism is proposed to update prototypes within a controllable range.

Finally, we adopt a guidance loss to measure the similarity between the current prototypes and the latent features, thereby optimizing the clustering structure of the learned representation. For the first view, since the knowledge space is empty, we directly use the prototypes $\mathbf{P}^v$ obtained through cross-attention to guide $\mathbf{H}^v$. For subsequent views, we use the prototypes $\mathbf{U}^v$ to adjust the distribution of $\mathbf{Z}^v$, and employ the prototypes $\mathbf{P}^B$ stored in the knowledge space to further guide $\mathbf{H}^v$. This design enables the features to benefit from both historical and current prototypes, thereby enhancing clustering performance. At the final clustering stage, we project the features of each newly arriving view into the null space of the preceding views and then concatenate the projected results. This ensures that previously learned features remain unaffected by new ones in the clustering space, thereby preventing degradation in clustering performance.

3.2.2. The Object Functions

For our task, the overall objective function can be formulated as:

$$\mathcal{L} = \beta\mathcal{L}_r + \gamma\mathcal{L}_g, \tag{18}$$

where $\mathcal{L}_r$ denotes the knowledge aggregation loss, $\mathcal{L}_g$ denotes the knowledge guidance loss, and β and γ are manually designed hyperparameters.

Since the collected information often contains substantial redundancy, we aim to extract feature representations that are beneficial for clustering. When the first view reaches , we reconstruct the latent features extracted by the autoencoder to obtain reconstruction features. Accordingly, the reconstruction loss is defined as:

$$\mathcal{L}_r = \|\mathbf{X}^v - \mathcal{F}_{\Phi^v}^v(\mathcal{E}_{\theta^v}^v(\mathbf{X}^v))\|_2^2, \tag{19}$$

where $\mathcal{E}_{\theta^v}^v(\cdot)$ and $\mathcal{F}_{\Phi^v}^v(\cdot)$ denote the encoder and decoder functions for the v-th view, parameterized by θ^v and Φ^v, respectively.

To optimize the feature distribution and enhance the clustering structure, we leverage the knowledge prototype to guide the current view's data distribution. Specifically, we first compute the similarity between each feature and the prototype knowledge:

$$M(\mathbf{h}_i^v, \mathbf{p}_j^v) = \exp(\mathrm{S}(\mathbf{h}_i^v, \mathbf{p}_j^v)), \tag{20}$$

where $\mathbf{h}_i^v$ denotes the i-th feature of the current view, and $\mathbf{p}_j^v$ denotes the j-th knowledge space prototype, S represents the similarity of $\mathbf{h}_i^v$ and $\mathbf{p}_j^v$.

Thus, the probability that the j-th prototype contains the i-th feature is given by:

$$G_{i,j} = \frac{\exp(\mathrm{S}(\mathbf{h}_i^v, \mathbf{p}_j^v))}{\sum_{k=1}^K \exp(\mathrm{S}(\mathbf{h}_i^v, \mathbf{p}_k^v))}, \tag{21}$$

where K is the total number of prototypes.

Finally, to optimize the clustering structure of the features, the knowledge-guided learning loss is defined as:

$$\mathcal{L}_g = \frac{1}{KN}\sum_{j=1}^K\sum_{i=1}^N \log G_{i,j}, \tag{22}$$

where N denotes the total number of features. Afterward, the prototype knowledge of the current view is uploaded to the knowledge space to guide the training of subsequently collected views.

4. Experiments

4.1. Datasets and Experiment Settings

We validate the effectiveness of ALMC by comparing it with other methods on six widely used datasets. Concretely, **ALOI-10** [9] contains 1,079 samples across 10 categories,

with each sample having four features: HSB, RGB, Colorsim, and Haralick. **ALOI-100** [9] contains 10,800 samples across 100 categories, with each sample having the same four features: HSB, RGB, Colorsim, and Haralick. **Outdoorscene** [26] contains 2,688 samples across 8 categories, with each sample represented by four types of features: GIST, HOG, SIFT, and Color histogram. **Oxford4** [27] contains 800 samples across 4 categories, with each sample represented by four feature types: SIFT, Color, GIST, and HOG. **NUS** [4] contains 2,400 samples across 31 categories, with each sample represented by four features: GIST, SIFT, Color, and HOG. **Aloideep3v** [9] contains 10,800 samples across 1,000 categories, with each sample represented by three deep features extracted from different neural networks.

Our ALMC model is built using PyTorch 2.3.0, and all experiments are conducted on a Linux machine equipped with an NVIDIA GPU and 32GB of RAM. For training, each acquired view is run for 200 epochs with a batch size of 256 and a learning rate of 0.0001. The Adam optimizer is applied for parameter updates, and ReLU serves as the activation function throughout the network. Please see the **Appendix** for instructions on setting hyperparameters.

4.2. Comparison Experiments

We conduct comparative experiments to evaluate ALMC against state-of-the-art methods, including eight DMVC methods and two SVC methods. Specifically, these methods are: ProImp [13], DCP [21], APADC [46], CPSPAN [11], ICMVC [2], MAGA [1], FMCSC [45], MCMVC [8], LSVC [14], and DSVC [53].

We evaluate clustering performance using Accuracy (ACC), Normalized Mutual Information (NMI), and Adjusted Rand Index (ARI). ACC measures the proportion of correctly clustered samples, NMI quantifies the mutual dependence between predicted and true labels, and ARI assesses clustering agreement while correcting for chance.

4.2.1. Comparison Results Analysis

As shown in Tab. 1 and Tab. 2, we compare the proposed ALMC method with several state-of-the-art multi-view clustering approaches on six benchmark datasets (Oxford4, ALOI-10, NUS, OutdoorScene, ALOI-100, and Aloideep3v). The results indicate that our method achieves the best performance on five datasets and the second-best performance on the NUS dataset.

Specifically, on the Oxford4 and ALOI-10 datasets, ALMC achieves the best performance across all three evaluation metrics (*e.g.*, ACC, NMI, and ARI). In particular, on the ALOI-10 dataset, ALMC attains an ACC of 90.90%, significantly outperforming the second-best methods ICMVC (87.38%) and DSVC (85.43%), which strongly validates the effectiveness of our core mechanism *e.g.*, updating the knowledge space by projecting new knowledge into the null space of the old subspace. This mechanism enables ALMC to learn new view information while maximally preserving previously acquired knowledge, thereby avoiding the common performance degradation or knowledge degradation observed in conventional lifelong learning models. Moreover, the relatively low variance across all metrics indicates that ALMC is highly robust and insensitive to data stream uncertainty.

As shown in Tab. 2, ALMC consistently outperforms existing methods on larger and more complex datasets. On OutdoorScene, it achieves clear advantages across all metrics, reflecting strong generalization in challenging visual scenarios. Similarly, on ALOI-100, with 100 categories, ALMC attains the highest ACC, NMI, and ARI, demonstrating that its continuously evolving knowledge space remains robust as task difficulty and data complexity increase. Compared with DSVC, a representative streaming-view clustering method, ALMC avoids the information loss caused by distribution consistency alignment and fully leverages knowledge from all views rather than relying solely on current-view prototypes. Overall, through its lifelong anti-degradation mechanism , ALMC achieves consistently high performance across diverse streaming datasets, effectively mitigating knowledge degradation and ensuring stable lifelong multi-view clustering.

4.2.2. Visualization

To more intuitively demonstrate the effectiveness of ALMC, we conduct visualization experiments on the ALOI-10 dataset. As shown in Fig. 3, in dynamic tasks, the boundaries between clusters become increasingly clear and the intra-cluster structures more compact as additional views are accumulated. As illustrated in Fig. 4, compared with both streaming-view clustering methods and static DMVC frameworks, ALMC exhibits superior inter-cluster separability and intra-cluster compactness. Moreover, the clustering results of our method are more cohesive: for example, in DSVC, samples of the same category are split into two clusters (pink), and in MCMVC, some category samples are even mixed with other categories (pink). This demonstrates the effectiveness of guiding feature clustering through knowledge space prototypes in ALMC, and confirms that uploading new knowledge via null-space projection preserves existing knowledge in the knowledge space while incorporating new information. The null-space projection and concatenation in the clustering space also ensure that existing features remain unaffected by newly added features. Together, these factors contribute to the superior accuracy and reliability of the clustering results.

4.2.3. View Sensitivity Analysis

To comprehensively evaluate the performance of ALMC in dynamic tasks, we shuffled the order of views to verify the superiority and stability of our method under different

Table 1. Comparative results of different methods on Oxford4, ALOI-10 and NUS dataset. Each value is reported as mean ± variance. Red indicates the best value, blue indicates the second best.

Method	Oxford4			ALOI-10			NUS		
	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
ProImp (IJCAI'23)	33.40 ± 2.57	2.37 ± 0.96	1.85 ± 1.05	84.34 ± 3.37	81.86 ± 2.56	74.96 ± 3.28	21.21 ± 0.29	9.43 ± 0.23	4.60 ± 0.16
DCP (TPAMI'23)	30.73 ± 2.36	2.88 ± 1.05	1.62 ± 1.15	56.31 ± 4.30	63.39 ± 4.57	40.68 ± 4.43	19.24 ± 0.72	7.59 ± 0.47	2.76 ± 0.58
APADC (TIP'23)	32.25 ± 1.22	3.40 ± 0.52	2.75 ± 0.34	67.84 ± 2.77	68.43 ± 3.21	55.48 ± 1.89	21.92 ± 0.50	9.26 ± 0.30	4.82 ± 0.94
CPSPAN (CVPR'23)	36.01 ± 1.05	5.04 ± 0.51	4.15 ± 0.27	64.45 ± 4.90	77.01 ± 1.33	57.25 ± 4.56	21.35 ± 1.61	9.67 ± 0.84	4.43 ± 0.75
ICMVC (AAAI'24)	31.31 ± 1.15	1.79 ± 0.32	1.18 ± 0.29	87.38 ± 7.28	85.27 ± 4.51	79.54 ± 8.36	20.76 ± 0.71	9.44 ± 0.24	4.47 ± 0.20
MAGA (IF'24)	33.17 ± 0.79	4.08 ± 0.64	3.27 ± 0.71	57.44 ± 8.40	62.64 ± 6.61	43.57 ± 8.50	25.86 ± 0.53	14.94 ± 0.04	7.99 ± 0.40
FMCSC (NeurIPS'24)	33.75 ± 1.91	4.47 ± 0.67	3.23 ± 0.54	85.17 ± 3.82	83.98 ± 1.05	77.53 ± 2.29	23.88 ± 1.22	11.43 ± 0.48	6.53 ± 0.58
MCMVC (TPAMI'24)	30.95 ± 1.90	2.03 ± 0.51	1.29 ± 0.61	55.96 ± 6.23	58.76 ± 4.08	43.75 ± 5.07	17.67 ± 0.97	8.29 ± 0.93	1.59 ± 0.19
LSVC (TNNLS'24)	33.12 ± 1.78	3.26 ± 0.82	2.21 ± 0.51	66.90 ± 4.01	67.81 ± 3.43	60.53 ± 4.12	19.62 ± 0.84	8.06 ± 0.95	4.17 ± 0.25
DSVC (ICML'25)	36.30 ± 1.25	5.51 ± 1.59	4.63 ± 1.15	85.43 ± 5.39	84.77 ± 2.40	78.74 ± 4.27	24.15 ± 0.79	12.63 ± 0.46	6.44 ± 0.40
ALMC(Ours)	36.87 ± 0.97	5.79 ± 1.32	4.89 ± 0.94	90.90 ± 4.14	87.32 ± 1.93	83.37 ± 3.32	25.37 ± 1.09	13.46 ± 0.73	7.20 ± 0.67

Table 2. Comparative results of different methods on OutdoorScene, ALOI-100 and Aloideepv3 dataset. Each value is reported as mean ± variance. Red indicates the best value, blue indicates the second best.

Method	OutdoorScene			ALOI-100			Aloideep3v		
	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
ProImp (IJCAI'23)	50.15 ± 1.74	40.15 ± 1.01	28.95 ± 1.12	70.84 ± 1.67	82.71 ± 0.82	64.35 ± 2.08	77.60 ± 0.83	94.52 ± 0.52	78.10 ± 1.64
DCP (TPAMI'23)	48.60 ± 3.70	41.82 ± 2.68	22.2226 ± 3.22	60.42 ± 0.52	81.98 ± 0.61	51.12 ± 0.49	29.92 ± 1.22	75.43 ± 1.39	22.01 ± 1.09
APADC (TIP'23)	61.79 ± 4.66	48.48 ± 3.26	37.01 ± 3.01	36.38 ± 2.21	60.90 ± 4.16	24.86 ± 1.98	68.92 ± 2.68	90.72 ± 4.89	65.84 ± 3.11
CPSPAN (CVPR'23)	58.48 ± 3.52	49.86 ± 0.82	37.75 ± 1.36	63.70 ± 1.75	81.86 ± 1.23	52.42 ± 3.50	76.88 ± 6.76	93.24 ± 1.67	76.12 ± 5.54
ICMVC (AAAI'24)	44.81 ± 3.22	25.75 ± 2.43	19.15 ± 2.64	58.45 ± 18.36	79.96 ± 6.81	44.70 ± 17.52	52.68 ± 7.85	83.79 ± 6.01	35.92 ± 3.30
MAGA (IF'24)	34.86 ± 1.50	22.69 ± 3.03	14.11± 2.12	52.98 ± 0.91	69.57 ± 0.57	40.45 ± 0.48	59.70 ± 0.07	82.88 ± 3.46	53.16 ± 7.35
FMCSC (NeurIPS'24)	56.80 ± 5.45	46.49 ± 3.51	37.29 ± 4.20	57.06 ± 1.70	65.90 ± 0.64	57.46 ± 0.12	88.17 ± 0.17	95.83 ± 0.32	87.19 ± 0.34
MCMVC (TPAMI'24)	49.98 ± 0.77	44.21 ± 1.84	27.20 ± 4.95	20.58 ± 1.46	45.67 ± 0.72	11.19 ± 0.72	37.04 ± 1.71	79.63 ± 1.32	37.21 ± 6.41
LSVC (TNNLS'24)	60.01 ± 4.52	48.75 ± 3.96	38.50 ± 4.51	64.11 ± 0.66	74.15 ± 0.48	55.01 ± 1.49	O/M	O/M	O/M
DSVC (ICML'25)	58.48 ± 2.63	49.87 ± 1.67	40.51 ± 1.72	80.53 ± 0.59	89.01 ± 0.18	73.96 ± 0.82	86.28 ± 1.54	96.36 ± 0.28	85.97 ± 1.34
ALMC(Ours)	64.26 ± 4.85	50.85 ± 1.86	43.03 ± 2.50	82.62 ± 1.28	89.74 ± 0.39	75.72 ± 0.97	88.18 ± 0.75	96.53 ± 0.44	87.21 ± 1.04

view sequences. Specifically, as shown in Fig. 5, with the continual accumulation of new views, ALMC consistently demonstrates stable performance. Guided by knowledge and the ongoing accumulation of knowledge in the knowledge space, the precision steadily improves as more views are added. Moreover, the differences in precision across different view orders are minimal, highlighting the robustness of our method under varying view sequences.

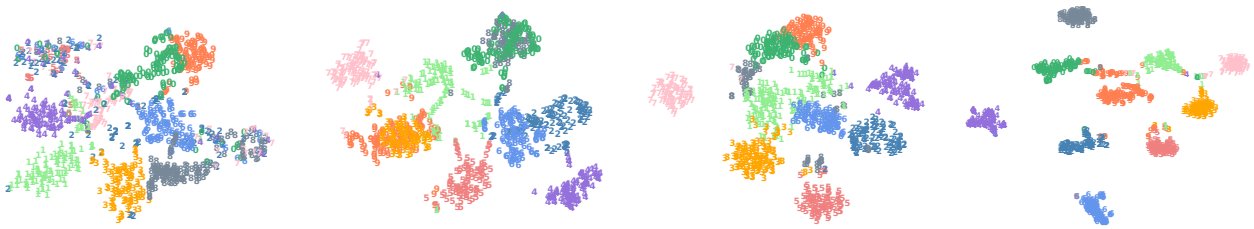
4.2.4. Parameter Sensitivity & Clustering Analysis

In ALMC, two hyperparameters, β and γ, are introduced to regulate the relative contributions of the knowledge aggregation and knowledge-guided learning losses. To assess the sensitivity of the model to these coefficients, their magnitudes are systematically adjusted within the interval of 10^2 to 10^{-4}. As depicted in Fig. 6a, the results indicate that both excessively high and extremely low values deteriorate clustering accuracy, underscoring the importance of achieving an appropriate balance among the three loss terms. Empirically, the best overall performance is obtained when both β and γ are chosen within the range of 10^1 to 10^{-1}.

In Fig. 6b, we evaluate the clustering performance of LSVC and ALMC on the OutdoorScene dataset. It is clear that clustering performance of LSVC becomes highly unstable with the increase in views, while clustering performance of ALMC remains very stable and steadily increases. This indicates that knowledge-space-guided feature clustering performs more effectively with increasing knowledge.

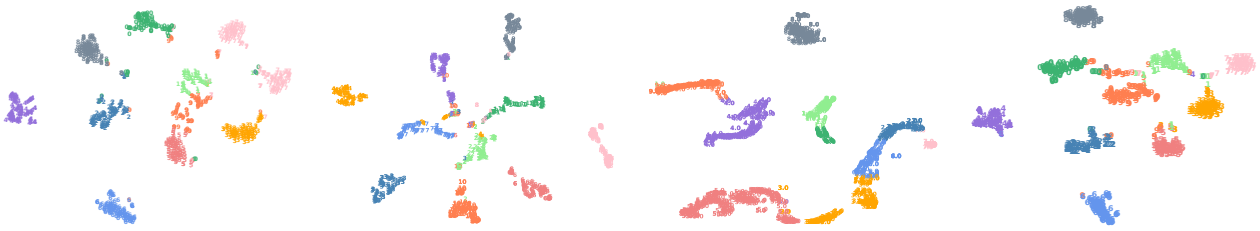
4.3. Ablation Study

Our ALMC framework consists of three components: the Knowledge Aggregation Learning module, the Knowledge Space, and the Knowledge-Guided Learning module. To evaluate the effectiveness of each component, we conducted ablation experiments on four variants of ALMC across six different datasets. As shown in Tab. 3, we remove key loss functions along with their corresponding network components. It is worth noting that we replace the null-space projection with alignment of old and new knowledge through a distribution consistency loss. In detail, Model I removes both L_g and the null-space projection, Model II removes only the null-space projection, Model III removes L_r, and Model IV removes L_g, while using distribution consistency to replace the null-space projection. The removal of any single component leads to a drop in performance. ALMC achieves optimal performance only when all three loss functions are included. The results indicate that the L_r loss preserves representative features of the data, maintaining data integrity and accuracy, while the L_g loss leverages prototype knowledge to guide the current view's data distribu-



(a) 1-st view (b) 2-nd view (c) 3-rd view (d) 4-th view

Figure 3. Visualization results of ALMC on the ALOI-10 dataset as the number of streaming views increases.

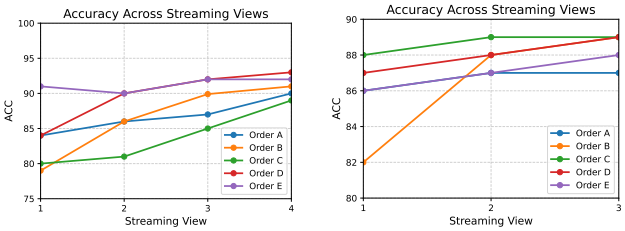


(a) DSVC (b) ICMVC (c) MCMVC (d) ALMC(Ours)

Figure 4. t-SNE visualization of the representations learned by DSVC, ICMVC, MCMVC, and our ALMC on the ALOI-10 dataset.

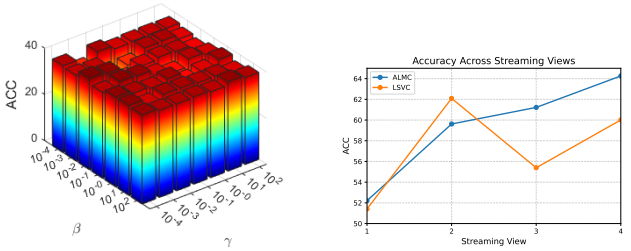
Table 3. Ablation results of different methods. Each value is reported as mean. Red indicates the best value, blue indicates the second best.

Datasets	Oxford4		ALOI-10		NUS		OutdoorScene		ALOI-100		Aloideep3v	
	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI
Model I	31.45	4.60	52.81	63.70	17.44	8.00	40.51	36.89	61.98	78.93	77.40	92.62
Model II	33.75	4.19	76.92	80.08	24.46	13.45	55.69	50.09	77.33	87.16	87.91	95.82
Model III	35.88	4.27	81.28	82.37	22.54	10.65	56.29	45.71	60.95	76.65	72.19	88.25
Model IV	34.88	4.24	82.21	83.96	24.46	12.45	55.80	50.41	81.31	88.99	86.04	96.35
Full Model	36.87	5.79	90.90	87.32	25.37	13.46	64.26	50.85	82.62	89.74	88.18	96.53



(a) ALOI-10 (b) Aloideepv3

Figure 5. The clustering performance with different sequences of views streaming on two datasets.



(a) Parameter sensitivity analysis on the Oxford4 dataset (b) Performance comparison between LSVC and ALMC

Figure 6. Parameter analysis and performance comparison.

tion, thereby enhancing the clustering structure.

5. Conclusion

This work proposes an anti-degradation lifelong learning paradigm to perform streaming multi-view clustering. By projecting new knowledge into the null space of existing knowledge, ALMC prevents performance degradation and ensures algorithmic stability. Combined with orthogonal prototype updates, EMA refinement, and prototype-guided alignment, it achieves continuous and robust representation learning. Extensive experiments and theoretical analysis validate the effectiveness and clear superiority of ALMC over existing methods.

6. Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant no. 62576298), the Sichuan Science and Technology Program (Grant nos. 2026NSFSC0453), the Mianyang Science and Technology Program (Grant no. 2025ZYDF096), and the Outstanding Talent Cultivation and Introduction Program (Grant no. 25ZX7182).

References

[1] Jintang Bian, Xiaohua Xie, Jian-Huang Lai, and Feiping Nie. Multi-view contrastive clustering via integrating graph aggregation and confidence enhancement. *Information Fusion*, 108:102393, 2024. 1, 2, 6

[2] 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. 6

[3] Xinyue Chen, Yazhou Ren, Jie Xu, Fangfei Lin, Xiaorong Pu, and Yang Yang. Bridging gaps: Federated multi-view clustering in heterogeneous hybrid views. *Advances in Neural Information Processing Systems*, 37:37020–37049, 2024. 1, 2

[4] Tat-Seng Chua, Jinhui Tang, Richang Hong, Haojie Li, Zhiping Luo, and Yantao Zheng. Nus-wide: a real-world web image database from national university of singapore. In *Proceedings of the ACM international conference on image and video retrieval*, pages 1–9, 2009. 6

[5] Zhibin Dong, Meng Liu, Siwei Wang, Ke Liang, Yi Zhang, Suyuan Liu, Jiaqi Jin, Xinwang Liu, and En Zhu. Enhanced then progressive fusion with view graph for multi-view clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 15518–15527, 2025. 1, 2

[6] Yu Feng, Weixuan Liang, Xinhang Wan, Jiyuan Liu, Suyuan Liu, Qian Qu, Renxiang Guan, Huiying Xu, and Xinwang Liu. Incremental nyström-based multiple kernel clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2025. 1, 2

[7] Yulu Fu, Yuting Li, Qiong Huang, Jinrong Cui, and Jie Wen. Anchor graph network for incomplete multiview clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 36(2):3708–3719, 2025. 1, 2

[8] Chuanxing Geng, Aiyang Han, and Songcan Chen. Explicit view-labels matter: A multifacet complementarity study of multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024. 6

[9] 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. 5, 6

[10] Ruiming Guo, Mouxing Yang, Yijie Lin, Xi Peng, and Peng Hu. Robust contrastive multi-view clustering against dual noisy correspondence. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 121401–121421, 2024. 2

[11] Jiaqi Jin, Siwei Wang, Zhibin Dong, Xinwang Liu, and En Zhu. Deep incomplete multi-view clustering with cross-view partial sample and prototype alignment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11600–11609, 2023. 6

[12] Zhao Kang, Wangtao Zhou, Zhitong Zhao, Junming Shao, Meng Han, and Zenglin Xu. Large-scale multi-view subspace clustering in linear time. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 4412–4419, 2020. 1

[13] Haobin Li, Yunfan Li, Mouxing Yang, Peng Hu, Dezhong Peng, and Xi Peng. Incomplete multi-view clustering via prototype-based imputation. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pages 3911–3919, 2023. 6

[14] Haoran Li, Zhenwen Ren, Yulan Guo, Jiali You, and Xiaojian You. Lsvc: A lifelong learning approach for stream-view clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 36(5):8714–8727, 2025. 2, 3, 6

[15] Xuelong Li, Han Zhang, Rong Wang, and Feiping Nie. Multiview clustering: A scalable and parameter-free bipartite graph fusion method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1):330–344, 2020. 1

[16] 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*, page 101941, 2023. 1

[17] Xingfeng Li, Yuangang Pan, Yinghui Sun, Quansen Sun, Ivor W Tsang, and Zhenwen Ren. Fast unpaired multi-view clustering. In *IJCAI*, pages 4488–4496, 2024. 1

[18] Xingfeng Li, Yuangang Pan, Yuan Sun, Yinghui Sun, Quansen Sun, Zhenwen Ren, and Ivor W Tsang. Scalable unpaired multi-view clustering with bipartite graph matching. *Information Fusion*, 116:102786, 2025. 1

[19] Xingfeng Li, Yuangang Pan 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*, pages 7087–7098, 2025. 1

[20] Jia-Qi Lin, Xiang-Long Li, Man-Sheng Chen, Chang-Dong Wang, and Haizhang Zhang. Incomplete data meets uncoupled case: A challenging task of multiview clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 2022. 1

[21] 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. 6

[22] Jiyuan Liu, Xinwang Liu, Xinhang Wan, Ke Liang, Weixuan Liang, Sihang Zhou, Huijun Wu, and Kehua Guo. Incomplete multi-view deep clustering with data imputation and alignment. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*. 2

[23] Jiyuan Liu, Xinwang Liu, Chuankun Li, Xinhang Wan, Hao Tan, Yi Zhang, Weixuan Liang, Qian Qu, Yu Feng, Renxiang Guan, and Ke Liang. Large-scale multi-view tensor

clustering with implicit linear kernels. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 20727–20736, 2025. 1, 2

[24] Jiyuan Liu, Xinwang Liu, Siqi Wang, Xinhang Wan, Dongsheng Li, Kai Lu, and Kunlun He. Communication-efficient federated multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025. 1, 2

[25] Suyuan Liu, Siwei Wang, Ke Liang, Junpu Zhang, Zhibin Dong, Tianrui Liu, En Zhu, Xinwang Liu, and Kunlun He. Alleviate anchor-shift: Explore blind spots with cross-view reconstruction for incomplete multi-view clustering. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 87509–87531, 2024. 1, 2

[26] Srinivasa G Narasimhan, Chi Wang, and Shree K Nayar. All the images of an outdoor scene. In *European conference on computer vision*, pages 148–162. Springer, 2002. 6

[27] Maria-Elena Nilsback and Andrew Zisserman. Automated flower classification over a large number of classes. In *2008 Sixth Indian conference on computer vision, graphics & image processing*, pages 722–729. IEEE, 2008. 6

[28] Yazhou Ren, Xinyue Chen, Jie Xu, Jingyu Pu, Yonghao Huang, Xiaorong Pu, Ce Zhu, Xiaofeng Zhu, Zhifeng Hao, and Lifang He. A novel federated multi-view clustering method for unaligned and incomplete data fusion. *Information Fusion*, 108:102357, 2024. 1

[29] Nina Shvetsova, Brian Chen, Andrew Rouditchenko, Samuel Thomas, Brian Kingsbury, Rogerio S. Feris, David Harwath, James Glass, and Hilde Kuehne. Everything at once – multimodal fusion transformer for video retrieval. In *CVPR*, pages 20020–20029, 2022. 2

[30] Gan Sun, Yang Cong, Qianqian Wang, Jun Li, and Yun Fu. Lifelong spectral clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 5867–5874, 2020. 3

[31] Gan Sun, Yang Cong, Jiahua Dong, Yuyang Liu, Zhengming Ding, and Haibin Yu. What and how: Generalized lifelong spectral clustering via dual memory. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7): 3895–3908, 2022. 3

[32] 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. 1, 2

[33] Yuan Sun, Yongxiang Li, Zhenwen Ren, Guiduo Duan, Dezhong Peng, and Peng Hu. Roll: Robust noisy pseudo-label learning for multi-view clustering with noisy correspondence. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 30732–30741, 2025. 2

[34] Xinhang Wan, Jiyuan Liu, Weixuan Liang, Xinwang Liu, Yi Wen, and En Zhu. Continual multi-view clustering. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 3676–3684, 2022. 3

[35] Xinhang Wan, Bin Xiao, Xinwang Liu, Jiyuan Liu, Weixuan Liang, and En Zhu. Fast continual multi-view clustering with incomplete views. *IEEE Transactions on Image Processing*, 33:2995–3008, 2024. 3

[36] Xinhang Wan, Jiyuan Liu, Qian Qu, Suyuan Liu, Chuyu Zhang, Fangdi Wang, Xinwang Liu, En Zhu, and Kunlun He. Intra-view and inter-view correlation guided multi-view novel class discovery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025. 2

[37] Xinhang Wan, Jiyuan Liu, Hao Yu, Qian Qu, Ao Li, Xinwang Liu, Ke Liang, Zhibin Dong, and En Zhu. Contrastive continual multiview clustering with filtered structural fusion. *IEEE Transactions on Neural Networks and Learning Systems*, 36(8):14829–14842, 2025. 3

[38] Changwei Wang, Rongtao Xu, Ke Lu, Shibiao Xu, Weiliang Meng, Yuyang Zhang, Bin Fan, and Xiaopeng Zhang. Attention weighted local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9):10632–10649, 2023. 1

[39] Fangdi Wang, Jiaqi Jin, Jingtao Hu, Suyuan Liu, Xihong Yang, Siwei Wang, Xinwang Liu, and En Zhu. Evaluate then cooperate: Shapley-based view cooperation enhancement for multi-view clustering. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 135355–135379, 2024. 1, 2

[40] Jing Wang, Songhe Feng, Kristoffer Knutsen Wickstrøm, and Michael C. Kampffmeyer. Adaptcmvc: Robust adaption to incremental views in continual multi-view clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10285–10294, 2025. 3

[41] Shuqin Wang, Yongyong Chen, Shuang Yi, and Guoqing Chao. Frobenius norm-regularized robust graph learning for multi-view subspace clustering. *Applied Intelligence*, 52 (13):14935–14948, 2022. 1

[42] Jie Wen, Chengliang Liu, Shanshan Deng, Waikeung Wong, Yong Xu, Chao Huang, Lunke Fei, and Guoqing Chao. Highly confident local structure based consensus graph learning for incomplete multi-view clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 15712–15721, 2023. 1, 2

[43] Jie Wen, Shijie Deng, Waikeung Wong, Guoqing Chao, Chao Huang, Lunke Fei, and Yong Xu. Diffusion-based missing-view generation with the application on incomplete multi-view clustering. In *Proceedings of the 41st International Conference on Machine Learning (ICML), PMLR*, pages 52762–52778, 2024. 2

[44] Jingqian Wu, Rongtao Xu, Zach Wood-Doughty, Changwei Wang, Shibiao Xu, and Edmund Y Lam. Segment anything model is a good teacher for local feature learning. *IEEE Transactions on Image Processing*, 2025. 1

[45] Jie Xu Fangfei Lin Xiaorong Pu Yang Yang Xinyue Chen, Yazhou Ren. Bridging gaps: Federated multi-view clustering in heterogeneous hybrid views. In *NeurIPS*, pages 1–23, 2024. 6

[46] 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 Trans. Image Process.*, 32:1354–1366, 2023. 6

[47] Wanqi Yang, Like Xin, Lei Wang, Ming Yang, Wenzhu Yan, and Yang Gao. Iterative multiview subspace learning for un-

paired multiview clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 2023. 1

[48] Hongwei Yin, Wenjun Hu, Zhao Zhang, Jungang Lou, and Minmin Miao. Incremental multi-view spectral clustering with sparse and connected graph learning. *Neural Networks*, 144:260–270, 2021. 3

[49] Hong Yu, Jia Tang, Guoyin Wang, and Xinbo Gao. A novel multi-view clustering method for unknown mapping relationships between cross-view samples. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 2075–2083, 2021. 1

[50] Shengju Yu, Zhibin Dong, Siwei Wang, Suyuan Liu, Ke Liang, Xinwang Liu, and En Zhu. From spectrum-free towards baseline-view-free: Double-track proximity driven multi-view clustering. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2025. 1, 2

[51] Shengju Yu, Siwei Wang, Xinwang Liu, Yiu-ming Cheung, and En Zhu. Bifurcate then alienate: Incomplete multi-view clustering via coupled distribution learning with linear overhead. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2025. 2

[52] 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. 1, 2

[53] Honglin Yuan, Xingfeng Li, Jian Dai, Xiaojian You, Yuan Sun, and Zhenwen Ren. Deep streaming view clustering. In *Forty-second International Conference on Machine Learning*, 2025. 2, 3, 6

[54] 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. 1, 2

[55] Chao Zhang, Deng Xu, Xiuyi Jia, Chunlin Chen, and Huaxiong Li. Continual multi-view clustering with consistent anchor guidance. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)*, pages 5434–5442, 2024. 3

[56] Xinyue Zhang, Zhenwen Ren, and Chao Yang. Center consistency guided multi-view embedding anchor learning for large-scale graph clustering. *Knowledge-Based Systems*, 260:110162, 2023. 1

[57] Yi Zhang, Siwei Wang, Jiyuan Liu, Shengju Yu, Zhibin Dong, Suyuan Liu, Xinwang Liu, and En Zhu. Dleft-mkc: Dynamic late fusion multiple kernel clustering with robust tensor learning via min-max optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2025. 1, 2

[58] Peng Zhou, Yi-Dong Shen, Liang Du, Fan Ye, and Xuejun Li. Incremental multi-view spectral clustering. *Knowledge-Based Systems*, 174:73–86, 2019. 3