CONFIDENCE-WEIGHTED DUAL-EXPERT CORRESPONDENCE REFINEMENT FOR INCOMPLETE MULTI-VIEW CLUSTERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Incomplete multi-view clustering (IMVC) aims to partition all samples effectively when each sample is observed in only a subset of views. However, most existing IMVC methods commonly treat the observed cross-view correspondences as equally reliable supervision, while overlooking the varying semantic and structural consistency of these correspondences under incomplete and heterogeneous observations. To address this limitation, we propose an anchored dual-expert framework that progressively learns confidence-weighted cross-view correspondences from complementary first- and second-order structural information. Specifically, we first exploit the observed partial overlap to initialize soft cross-view correspondences and select an informative anchor view, providing an initial basis for correspondence refinement. Furthermore, a dual-expert matching mechanism is developed, which combines a Koopmans–Beckmann QAP expert for modeling first-order node and local relational affinities with an anchor-based Lawler QAP expert for capturing higher-order structural consistency in a compact learned anchor space. The latter incorporates an outlier state to accommodate unmatched structures caused by incomplete observations. Finally, we introduce a momentum-based consistency refinement mechanism to fuse the complementary evidence from the two experts and iteratively update correspondence confidences, which in turn provide progressively refined supervision for multi-view representation learning. The refined cross-view and within-view affinities are subsequently integrated to construct a global affinity matrix for spectral clustering. Extensive experiments under varying view missing rates demonstrate the effectiveness and robustness of the proposed framework.

1 INTRODUCTION

Multi-view clustering seeks to integrate data from multiple views and exploit their complementarity and consistency, thereby improving clustering accuracy and robustness (Fang et al., 2023; Zeng et al., 2024; Guo et al., 2024; Zhao et al., 2023). However, in real-world applications, an object is often observed in only a subset of views due to sensor failures, privacy constraints, or the high cost of data collection. Such view missingness violates the assumption that all views are fully observed in conventional MVC, leaving only limited cross-view correspondences and making effective integration of multi-view information more challenging. Incomplete multi-view clustering (IMVC)(Yu et al., 2025; Xu et al., 2024; Wen et al., 2024) addresses this realistic setting, where each view contains only a subset of samples and the sets of observed samples may differ across views.

Existing IMVC methods broadly follow two paradigms: imputation-based and imputation-free. Imputation-based methods first reconstruct missing views and then perform multi-view clustering. For example, COMPLETER (Lin et al., 2021) employs view-specific autoencoders to reconstruct missing views and enforces consistency between observed and reconstructed representations. Dual contrastive prediction (Lin et al., 2023) further strengthens this paradigm by pulling the observed and predicted representations of the same sample together while separating negative pairs. In contrast, imputation-free methods bypass explicit view reconstruction and directly exploit the available observations for clustering. For example, CPSPAN (Jin et al., 2023) learns a shared latent space and consensus representation from the observed views to capture their complementary information,

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

while AGC-IMC (Wen et al., 2021) constructs and integrates view-specific affinity graphs to preserve complementary local structures for clustering. More recently, cross-view alignment methods establish correspondences between co-observed samples and align their representations to learn a consensus structure (Dai et al., 2025; Ding et al., 2025).

These two paradigms, however, face complementary limitations in handling missing views. Imputation-based methods reconstruct missing views before clustering, but reconstruction errors may propagate to subsequent representation learning and affinity construction. This motivates avoiding explicit imputation and instead exploiting the observations that are directly available. Imputation-free methods avoid reconstruction errors, but must effectively integrate incomplete and heterogeneous observations across views. Recent cross-view alignment methods exploit correspondences among co-observed samples to learn a consensus structure (Dai et al., 2025; Ding et al., 2025). However, sample identity alone provides correspondence supervision without revealing how informative each relation is for clustering. Under substantial view heterogeneity and missingness, some co-observed samples may exhibit clear cross-view structure, whereas others may yield ambiguous or noisy signals. Treating these correspondences uniformly can therefore weaken the learned correspondence and affinity structures. This motivates incorporating structural information to distinguish reliable correspondences from ambiguous ones. However, explicitly modeling second-order relations over all samples is computationally prohibitive, as dense second-order matching incurs $O(N^4)$ complexity, making such matching costly for large-scale data.

To address these challenges, we formulate incomplete multi-view clustering as confidence-weighted cross-view correspondence refinement, where reliable correspondences are progressively identified and refined using complementary structural cues. Based on this formulation, we propose an anchored dual-expert align–fuse–refine framework (Lin et al., 2026). An informative anchor view is first selected, and soft cross-view correspondences are initialized from observed overlaps via mutual top-K matching in a shared PCA space. We then refine these correspondences using two complementary QAP experts: KB-QAP captures first-order node and edge affinities, whereas Lawler QAP models second-order relations in a compact anchor space. A momentum-based cooperation mechanism exploits agreement between the two experts to reinforce reliable correspondences and gradually suppress uncertain ones. This design operates directly on observed data without imputing missing views, while anchor-space matching avoids the prohibitive cost of sample-level high-order matching. Finally, the refined view-specific and cross-view affinities are integrated into a global affinity matrix for spectral clustering. Our main contributions are as follows:

- We reformulate IMVC as confidence-weighted cross-view correspondence refinement, rather than treating all observed pairs equally. We further develop automatic anchor-view selection and mutual top-K bootstrapping to initialize soft correspondences.
- We extend the dual-expert QAP framework from supervised graph matching to unsupervised IMVC, modeling higher-order cross-view correspondences in a compact learnable anchor space and reducing the matching complexity from $O(N^4)$ to $O(M^4)$ with $M \ll N$.
- Experiments on seven benchmarks under four missing rates show our method achieves the best average ACC, NMI, and ARI among eight baselines.

2 Related Work

Depending on whether the missing views are explicitly recovered, existing IMVC methods can be broadly categorized into two groups: completion-based methods and imputation-free methods.

Completion-based methods explicitly reconstruct missing views before applying standard multi-view clustering pipelines. A common strategy is to exploit cross-view semantic or structural information to guide the recovery process. LRGR Xi et al. (2025) refines locally imputed representations with view-specific prototype guidance and subsequently realigns cross-view semantics through pseudo-label consistency. DIMVC-HIA Du et al. (2026) follows a coarse-to-fine hierarchical strategy, first estimating missing cluster assignments from cross-view contrastive similarity and then reconstructing missing features using intra-cluster statistics. IMC-MCL Yin et al. (2025) uses dual prediction to infer missing data by minimizing conditional entropy between views, while jointly learning latent, high-level, and semantic representations. SPORT Guo et al. (2026) decomposes multi-view prototypes into shared and view-specific components and combines global prototype matching with local neighborhood matching for missing-view recovery.

Imputation-free methods learn consensus representations directly from observed views without reconstructing missing data. Early approaches such as DCP Lin et al. (2023) project different views into a common latent space while jointly preserving instance-level and cluster-level correlations, enabling fusion without explicit feature recovery. More recent methods increasingly operate at higher semantic levels. CPMN Wang et al. (2025) considers the fully incomplete setting without cross-view instance correspondence and aligns high-level cluster prototypes through contrastive matching. GLC He et al. (2026) jointly models global semantic correlations and local semantic non-correlations to alleviate rare- and mis-paired issues caused by missing views. FreeCSL Dai et al. (2025) further removes both imputation and explicit instance alignment by discovering a shared semantic space through consensus-prototype-based contrastive clustering.

3 METHOD

We formulate incomplete multi-view clustering as a correspondence refinement problem. We first warm up the view encoders and select an informative anchor view to initialize soft cross-view correspondences from observed overlaps. Two complementary experts then refine the correspondences: KB-QAP captures first-order node and local relational affinities, while anchor-based L-QAP captures higher-order structural consistency. Their predictions are integrated with historical confidences through momentum-based refinement, feeding updated correspondences back into representation learning and matching. Finally, the refined correspondences, within-view affinities, and common representation are fused for spectral clustering.

3.1 NOTATION AND PROBLEM SETUP

Let $\mathcal{V} = \{1, \dots, V\}$ be the set of views and $[N] = \{1, \dots, N\}$ the set of entities. Each view v observes a subset $\mathcal{E}^{(v)} \subseteq [N]$ of entities, with features $X^{(v)} \in \mathbb{R}^{N_v \times d_v}$ where $N_v = |\mathcal{E}^{(v)}|$, and a local-to-global index map $g^{(v)} : [N_v] \to [N]$ that identifies the entity corresponding to each row. Given the partially observed features $\{X^{(v)}\}$ and the known correspondences $\{g^{(v)}\}$, the goal is to partition $[N]$ into C clusters without ground-truth labels.

3.2 ANCHOR-GUIDED SOFT CORRESPONDENCE INITIALIZATION

Reliable correspondence is difficult to obtain directly under high missing rates. We therefore first stabilize the view representations, select an informative anchor view, and bootstrap soft cross-view correspondences from observed overlaps.

Anchor view selection and pseudo-label initialization. To bootstrap the correspondences, we select one view to serve as the reference for all cross-view alignments. To this end, we score each view by its completeness comp_v, feature informativeness var_v, and the spectral gap gap_v of a k-nearest-neighbor graph built on the observed rows of $X^{(v)}$. After min–max normalization, these three criteria are combined by a weighted sum with non-negative weights α, β, and ζ that sum to one, and the view with the highest score is selected as the anchor $v^\star$.

For each non-anchor view $v \neq v^\star$, we construct a soft pseudo-candidate matrix $\mathbf{Y}^{(v)} \in [0, 1]$. Both views are standardized, projected to a common subspace, and row-normalized into $\hat{\mathbf{Z}}^{(v)}$; the projected similarity is $\mathbf{S}^{(v)} = \hat{\mathbf{Z}}^{(v^\star)}(\hat{\mathbf{Z}}^{(v)})^\top$. Let $\mathcal{C}_i^{(v)}$ be the top-r target columns of row i of $\mathbf{S}^{(v)}$ and let $\mathcal{R}_j^{(v)}$ be the top-r anchor rows of column j. With τ_v the τ-quantile of the similarities over entries that are observed on both the row and the column and $S_{\max}^{(v)}$ their maximum, the mutual top-r pseudo-label is

$$Y_{ij}^{(v)} = m_i^{(v^\star)}\, m_j^{(v)}\, \mathbf{1}\Big[j \in \mathcal{C}_i^{(v)} \wedge i \in \mathcal{R}_j^{(v)}\Big] \left(\frac{1}{2} + \frac{1}{2}\,\mathrm{clip}_{[0,1]}\left(\frac{S_{ij}^{(v)} - \tau_v}{S_{\max}^{(v)} - \tau_v}\right)\right). \tag{1}$$

We treat the resulting $\mathbf{Y}^{(v)}$ as a per-pair soft confidence matrix rather than a verified instance correspondence.

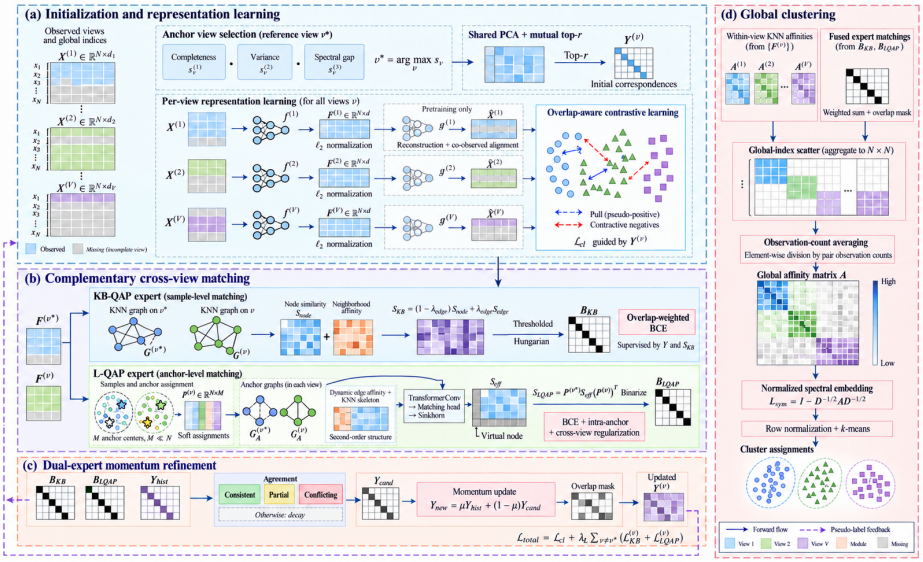

Figure 1: Our anchored dual-expert framework.

Overlap-aware contrastive representation learning. The initial pseudo-correspondences provide soft supervision for representation learning. Let $\mathbf{F}^{(v)}$ denote the ℓ_2-normalized representation produced by the encoder of view v. For the anchor view $v^\star$ and a target view v, we employ the cross-view InfoNCE loss

$$\mathcal{L}_{\mathrm{cl}}^{(v)} = -\frac{1}{|\mathcal{I}^{(v)}|} \sum_{i \in \mathcal{I}^{(v)}} \log \frac{\sum_{j \in \mathcal{P}_i^{(v)}} \exp\left(\mathbf{F}_{i:}^{(v^\star)} \cdot \mathbf{F}_{j:}^{(v)} / \tau_c\right)}{\sum_{j=1}^{N_v} \exp\left(\mathbf{F}_{i:}^{(v^\star)} \cdot \mathbf{F}_{j:}^{(v)} / \tau_c\right)}, \tag{2}$$

where $\tau_c > 0$ is a temperature. Crucially, the positive set uses the current pseudo-correspondence and is restricted to samples that are truly co-observed in both views. The total contrastive loss is $\mathcal{L}_{\mathrm{cl}} = \frac{1}{V-1} \sum_{v \neq v^\star} \mathcal{L}_{\mathrm{cl}}^{(v)}$.

3.3 Dual-Expert Matching Mechanism

The initialized correspondences are refined by two complementary matching experts that share the same supervision interface: the current confidence matrix $\mathbf{Y}^{(v)}$ serves as their soft target, and their similarity outputs are fused by the refinement module below.

First-order KB-QAP expert. Let $\mathbf{E}^{(v)}$ denote the unweighted adjacency matrix of the k-nearest-neighbor graph of view v. We compute node similarity $\mathbf{S}_{\mathrm{node}}$ from cross-view representation dot products and edge affinity $\mathbf{S}_{\mathrm{edge}}$ by propagating node similarities through the two neighborhood structures. Their weighted combination gives the Koopmans–Beckmann similarity $\mathbf{S}_{\mathrm{KB}}$. We optimize it against the soft correspondence target $\mathbf{Y}^{(v)}$ using an overlap-weighted binary cross-entropy:

$$\mathcal{L}_{\mathrm{BCE}}^{(v)} = -\frac{1}{|\Omega|} \sum_{(i,j) \in \Omega} w_{ij} \left[Y_{ij}^{(v)} \log \mathrm{sig}(\kappa S_{\mathrm{KB},ij}) + \left(1 - Y_{ij}^{(v)}\right) \log\left(1 - \mathrm{sig}(\kappa S_{\mathrm{KB},ij})\right) \right], \tag{3}$$

where Ω contains valid anchor–target pairs, $\kappa > 0$ is a temperature, and w_{ij} assigns larger weight to co-observed pairs. The overall KB-QAP loss $\mathcal{L}_{\mathrm{KB}}^{(v)}$ therefore comprises $\mathcal{L}_{\mathrm{BCE}}^{(v)}$, an intra-view structure consistency term, and a cross-view symmetry regularizer.

Second-order anchor-based L-QAP expert. Direct Lawler QAP incurs $O(N^4)$ complexity. We therefore perform matching in a compact anchor space with $M \ll N_{v^\star}$. K-means over the anchor-view features provides M centers, and each sample is softly assigned to these centers using a Gaussian kernel, producing normalized anchor-level representations $\mathbf{F}_{\mathrm{anc}}^{(v)}$.

We construct a second-order affinity over the M^2 directed anchor edges by mixing a Kronecker KNN skeleton prior with a feature-induced dynamic affinity, where the mixing weight decays linearly from 1 to 0.1 over training epochs. Two view-specific TransformerConv branches refine the anchor representations, whose inner products produce the matching logits $\mathbf{L}$. To account for unmatched structures, we augment $\mathbf{L}$ with a virtual outlier node and apply Sinkhorn normalization:

$$\widetilde{\mathbf{S}}_{\text{aug}} = \text{Sinkhorn}\left(\begin{bmatrix} \mathbf{L} & p\mathbf{1}_M \\ p\mathbf{1}_M^\top & b \end{bmatrix}\right), \qquad \widetilde{\mathbf{S}}_{\text{eff}} = [\widetilde{\mathbf{S}}_{\text{aug}}]_{1:M,1:M}. \tag{4}$$

The resulting anchor-level matching is projected back to the sample space to obtain $\mathbf{S}_{\text{LQAP}}$, while $\mathbf{Y}^{(v)}$ is projected into the anchor space to provide the corresponding soft target $\mathbf{Y}_{\text{anc}}$.

The L-QAP expert is optimized with a class-balanced binary cross-entropy on the augmented logits, together with an edge-consistency term and two structural regularizers: an intra-anchor transport term that minimizes the Frobenius distance between the Z-score-normalized anchor-space Gram matrix of the anchor view and that of the target view transported via $\mathbf{Y}_{\text{anc}}$, and a cross-view symmetry term that penalizes the asymmetry between the refined anchor affinity matrices of the two views. The edge-consistency term $\mathcal{L}_{\text{edge}}$ is defined as the negative quadratic form between the vectorized effective anchor assignment and the Lawler second-order affinity matrix, encouraging structurally consistent anchor matching. The overall L-QAP loss is

$$\mathcal{L}_{\text{LQAP}}^{(v)} = \mathcal{L}_{\text{BCE}} + \lambda_{\text{edge}}\,\mathcal{L}_{\text{edge}} + \lambda_{\text{intra}}\,\mathcal{L}_{\text{intra}} + \lambda_{\text{sym}}\,\mathcal{L}_{\text{sym}}. \tag{5}$$

Both experts thus share the same soft correspondence supervision while providing complementary first- and second-order matching evidence, $\mathbf{S}_{\text{KB}}$ and $\mathbf{S}_{\text{LQAP}}$, for the subsequent correspondence refinement.

3.4 Momentum-Based Consistency Refinement

Let $\mathbf{B}_{\text{KB}}^{(v)}$ and $\mathbf{B}_{\text{LQAP}}^{(v)}$ be binarized expert outputs and let $\mathbf{Y}_{\text{hist}}^{(v)}$ be the previous pseudo-labels. For each target view v, let $\Omega_{\text{ov}}^{(v)}$ be the co-observed block. We partition the pairs in $\Omega_{\text{ov}}^{(v)}$ into three regions: *consistent* pairs, where both experts and the history agree on a match; *partially consistent* pairs, where the history agrees with exactly one expert; and *conflicting* pairs, where both experts agree but the history does not. The candidate update is

$$Y_{\text{cand},ij}^{(v)} = \begin{cases} 1, & \text{consistent,} \\ (1-w_p)Y_{\text{hist},ij}^{(v)} + w_p/2, & \text{partially consistent,} \\ \text{clip}_{[0,1]}\left(\frac{1}{2}S_{\text{KB},ij} + \frac{1}{2}S_{\text{LQAP},ij}\right), & \text{conflicting,} \\ \delta_{ij}Y_{\text{hist},ij}^{(v)}, & \text{otherwise.} \end{cases} \tag{6}$$

The confidence matrix is then smoothed and masked to observed pairs:

$$\mathbf{Y}_{\text{new}}^{(v)} = \mu\,\mathbf{Y}_{\text{hist}}^{(v)} + (1-\mu)\,\mathbf{Y}_{\text{cand}}^{(v)}, \qquad \mathbf{Y}^{(v)} \leftarrow \mathbf{Y}_{\text{new}}^{(v)} \odot \left(\mathbf{m}^{(v^\star)}(\mathbf{m}^{(v)})^\top\right), \tag{7}$$

The updated matrix $\mathbf{Y}^{(v)}$ closes the supervision loop: in the next epoch it serves as the per-pair confidence target for both matching losses and, in turn, supervises the contrastive feature learning of the initialization module. After each refinement, both experts recompute their adaptive thresholds over the observed pairs.

3.5 Global Affinity and Spectral Clustering

At inference time, the refined correspondences and representations are integrated into a single global affinity. For each view v, a self-tuned local-scale KNN affinity is built from $\mathbf{F}^{(v)}$, giving the within-view affinity matrix. For each target view, the two expert similarities are fused under the overlap mask, giving the fused similarity between views $\mathbf{S}_{\text{fuse}}^{(v)}$. All within-view affinities and cross-view fused similarities are scattered into global matrices $\hat{\mathbf{A}}^{(v)}$ and $\hat{\mathbf{S}}^{(v)}$ using $g^{(v)}$:

$$A_{pq} = \frac{\sum_{v=1}^{V}\hat{A}_{pq}^{(v)} + \sum_{v\neq v^\star}\hat{S}_{pq}^{(v)}}{\max\left(\sum_{v=1}^{V}\mathbf{1}\left[p,q\in\mathcal{E}^{(v)}\right] + \sum_{v\neq v^\star}\mathbf{1}\left[p,q\in\mathcal{E}^{(v^\star)}\cap\mathcal{E}^{(v)}\right],\ 1\right)}. \tag{8}$$

In addition, a common representation $\mathbf{Z}_{\text{common}}$ is formed by averaging the observed normalized embeddings of each entity, a KNN affinity $\mathbf{A}_{\text{common}}$ is built from $\mathbf{Z}_{\text{common}}$, and the two affinities are blended. Finally, $\mathbf{A} \leftarrow (\mathbf{A}+\mathbf{A}^\top)/2$, and normalized spectral clustering computes the C smallest eigenvectors of $\mathbf{L}_{\text{sym}} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$, row-normalizes them, and applies k-means to obtain the cluster labels $\mathbf{z} \in [C]^N$.

3.6 Overall Training Objective

The full objective couples the contrastive loss of the initialization module with the two expert losses of the matching module for every target view,

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{cl}} + \lambda_L \sum\nolimits_{v \neq v^\star} \left(\mathcal{L}_{\text{KB}}^{(v)} + \mathcal{L}_{\text{LQAP}}^{(v)} \right), \qquad \lambda_L > 0. \tag{9}$$

During the main loop, all encoder and L-QAP module parameters are optimized jointly with Adam.

4 Experiments

4.1 Experimental Settings

Datasets and Missingness Protocol. We evaluated seven multi-view benchmarks that span different sample sizes, view counts, feature types, and cluster numbers, as summarized in Table 1.

Table 1: Dataset statistics used in the experiments.

Dataset	Samples	Views	Feature dimensions	Clusters
ALOI_100	10,800	4	$\{77, 13, 64, 125\}$	100
Hdigit	10,000	2	$\{784, 256\}$	10
Scene15	4,485	3	$\{20, 59, 40\}$	15
Wiki_fea	2,866	2	$\{128, 10\}$	10
HW	2,000	6	$\{216, 76, 64, 6, 240, 47\}$	10
100leaves	1,600	3	$\{64, 64, 64\}$	100
COIL20	1,440	3	$\{1024, 944, 4096\}$	20

Following the missing independent view-sample protocol, for a missing rate $r \in \{0.1, 0.3, 0.5, 0.7\}$, we independently mark each view of every sample as missing with probability r (equivalently, an expected fraction r of views are removed per sample). To avoid fully unobserved samples, if all views of a sample are marked missing, we randomly select one view and set it to observed. For each missing mask, every method uses the same set of initialization seeds across methods. We report the mean of ACC, NMI, and ARI over these runs.

Baselines. Eight baselines covering imputation-free, contrastive prediction, and subspace/graph learning IMVC models are compared: DCP (Lin et al., 2023), DSIMVC (Tang & Liu, 2022), DIMVC (Xu et al., 2022), RecFormer (Liu et al., 2024), FreeCSL (Dai et al., 2025), HSACC (Ding et al., 2025), DGIMVCM (Zhang et al., 2026), and the best single view (BSV) baseline. All baselines are evaluated under identical incomplete views/masks, following the official implementations and default settings of the respective papers when available.

Implementation Details. All experiments are implemented in PyTorch 2.4.1 and run on a single 48GB vGPU. Adam is used for optimization, with a linear warmup followed by cosine annealing. The learning rate, training epochs, and other hyperparameters are tuned separately for each dataset. The anchor-selection weights are fixed to $(0.6, 0.2, 0.2)$. For L-QAP, the number of learnable anchors is set to $M \in [40, 100]$, with 10–20 Sinkhorn iterations; its KNN skeleton is refreshed every s_{skel} epochs, where s_{skel} is tuned per dataset. Each view encoder is a two-layer MLP, and the L-QAP module uses a two-layer TransformerConv with four attention heads. The final affinity graph uses $k \in [4, 80]$ nearest neighbors, while the expert fusion weights are tuned per dataset.

4.2 Main Results

Following the above protocol, Table 2 summarizes the results on the seven multi-view benchmarks. For each dataset and missing rate $r \in \{0.1, 0.3, 0.5, 0.7\}$, we report the mean ACC, NMI, and ARI over five runs. More concretely, at the most challenging missing rate $r = 0.7$ on ALOI-100, our method outperforms FreeCSL by 8 percentage points in ACC. This indicates that confidence-weighted cross-view correspondence refinement is more robust to sparse observations than simply aligning all co-observed pairs uniformly, and that this advantage grows as the missing rate

increases. At the same time, by restricting second-order matching to a compact learnable anchor space, our method explicitly models cross-view correspondences even on 10k-sample benchmarks such as ALOI-100 and Hdigit, completes training without any view imputation, and substantially reduces the computational cost of correspondence refinement.

Table 2: Clustering results in ACC/NMI/ARI format. The best value of each metric is marked in bold red, and the second-best value of each metric is marked in blue.

Dataset	Method	r=0.1(ACC/NMI/ARI)	r=0.3(ACC/NMI/ARI)	r=0.5(ACC/NMI/ARI)	r=0.7(ACC/NMI/ARI)
100leaves	**Ours**	**0.9423/0.9616/0.8870**	**0.8331/0.8958/0.7291**	**0.6254/0.7824/0.4601**	**0.4381/0.6926/0.2687**
	FreeCSL	0.8544/0.9437/0.8113	0.6712/0.8285/0.5471	0.4381/0.6878/0.2746	0.3075/0.6052/0.1513
	DIMVC	0.6231/0.8147/0.4839	0.5049/0.7197/0.3191	0.3944/0.6403/0.2196	0.0825/0.2925/0.0139
	HSACC	0.4965/0.7035/0.2685	0.2391/0.5347/0.1014	0.1490/0.4347/0.0574	0.1241/0.3893/0.0087
	DGIMVCM	0.6594/0.8325/0.5175	0.5831/0.8010/0.4594	0.4325/0.7319/0.3074	0.3369/0.6640/0.1972
	RecFormer	0.7125/0.8509/0.5855	0.4944/0.7168/0.3045	0.3306/0.6224/0.1545	0.2944/0.6082/0.1340
	BSV	0.5169/0.7423/0.2653	0.4244/0.6389/0.0542	0.2994/0.5171/0.0124	0.2394/0.4355/0.0046
	DCP	0.5519/0.7946/0.4260	0.3835/0.6532/0.2839	0.3012/0.5997/0.1539	0.2294/0.5328/0.0831
	DSIMVC	0.3638/0.6872/0.2387	0.3081/0.6139/0.1774	0.2056/0.5067/0.0857	0.1781/0.4777/0.0661
HW	**Ours**	0.9515/0.8756/0.8872	0.9340/0.8694/0.8621	**0.9020**/0.8099/0.7978	**0.8640/0.7567/0.7352**
	FreeCSL	0.8650/0.8621/0.8113	0.8245/0.8382/0.7771	0.8170/0.7729/0.7168	0.7660/0.7300/0.6726
	DIMVC	0.4840/0.4702/0.3468	0.4170/0.4509/0.3062	0.3930/0.2871/0.1902	0.2480/0.2431/0.1212
	HSACC	0.7095/0.7342/0.5467	0.7375/0.7567/0.6483	0.6820/0.6804/0.5322	0.1770/0.1770/0.0240
	DGIMVCM	0.8345/0.8964/0.8247	0.9265/0.8762/0.8247	0.7650/0.7622/0.6772	0.3021/0.2846/0.1299
	RecFormer	**0.9645/0.9211/0.9225**	**0.9423/0.8875/0.8863**	0.8895/**0.8179/0.8058**	0.8110/0.7012/0.6322
	BSV	0.7180/0.6612/0.5698	0.6045/0.5549/0.3116	0.4330/0.3748/0.0904	0.3190/0.2964/0.0273
	DCP	0.8115/0.8393/0.7161	0.9010/0.8769/0.8598	0.6885/0.7329/0.6335	0.4160/0.4606/0.3139
	DSIMVC	0.8440/0.8112/0.7623	0.5190/0.4763/0.3198	0.5725/0.5479/0.4344	0.3775/0.3496/0.2029
ALOI_100	**Ours**	**0.9224**/0.9216/**0.8540**	**0.8719**/0.8776/0.7756	**0.8227**/0.8233/**0.6759**	**0.6192**/0.7061/**0.4497**
	FreeCSL	0.8650/**0.9408**/0.8278	0.8405/**0.9203/0.8021**	0.8033/**0.8810**/0.6726	0.5305/**0.7217**/0.4194
	DIMVC	0.4838/0.6995/0.3387	0.3438/0.5666/0.1205	0.4615/0.6119/0.2670	0.3640/0.5352/0.1643
	HSACC	0.6194/0.8084/0.5404	0.4744/0.6918/0.3496	0.2668/0.5410/0.1092	0.1450/0.2277/0.1315
	DGIMVCM	–/–/–	–/–/–	–/–/–	–/–/–
	RecFormer	0.6700/0.8225/0.5355	0.6212/0.7487/0.3881	0.4392/0.6304/0.2071	0.3105/0.5652/0.1497
	BSV	0.4152/0.6467/0.2163	0.3172/0.5396/0.0400	0.2357/0.4336/0.0089	0.1698/0.3238/0.0024
	DCP	0.5373/0.7581/0.3940	0.4291/0.6483/0.2928	0.3323/0.5572/0.2036	0.1876/0.4209/0.0765
	DSIMVC	–/–/–	–/–/–	–/–/–	–/–/–
scene15	**Ours**	0.4176/0.3988/0.2420	0.4198/0.3875/0.2357	**0.3634**/0.3148/0.1854	**0.3257/0.2745**/0.1459
	FreeCSL	0.3639/0.4081/0.2215	0.3761/0.3728/0.2174	0.3351/0.3044/0.1723	0.3178/0.1944/**0.1723**
	DIMVC	0.2803/0.3076/0.1391	0.3064/0.3189/0.1574	0.2491/0.2409/0.1097	0.2219/0.2224/0.0780
	HSACC	0.4047/0.4282/0.2391	0.3973/0.4030/0.2239	0.2883/0.2855/0.1110	0.1211/0.0542/0.0091
	DGIMVCM	**0.4896/0.4838/0.3116**	**0.4253/0.4128/0.2441**	0.3545/**0.3427**/0.1901	0.2185/0.2316/0.0900
	RecFormer	0.4013/0.3978/0.2207	0.3541/0.3569/0.1808	0.2990/0.2970/0.1377	0.2174/0.2217/0.0784
	BSV	0.3336/0.3295/0.1600	0.2769/0.2791/0.0791	0.2397/0.2254/0.0317	0.1989/0.1919/0.0128
	DCP	0.4054/0.4385/0.2475	0.3498/0.3582/0.1982	0.2948/0.3291/**0.1902**	0.1944/0.1891/0.0692
	DSIMVC	0.2845/0.3060/0.1468	0.2513/0.2690/0.1212	0.2526/0.2440/0.1151	0.2071/0.1736/0.0842
Hdigit	**Ours**	0.9544/0.9005/0.9128	0.9335/0.8528/0.8580	**0.8897/0.7735/0.7737**	**0.7592/0.7004/0.6431**
	FreeCSL	**0.9853/0.9598/0.9676**	**0.9538/0.8907/0.8997**	0.7415/0.6882/0.6119	0.6786/0.6258/0.5167
	DIMVC	0.3752/0.2849/0.1511	0.3587/0.2848/0.1510	0.2947/0.2722/0.1440	0.3399/0.2755/0.1297
	HSACC	0.9365/0.8702/0.8667	0.9013/0.7897/0.7943	0.8306/0.6856/0.6690	0.6764/0.5092/0.4342
	DGIMVCM	–/–/–	–/–/–	–/–/–	–/–/–
	RecFormer	0.7289/0.6439/0.5391	0.2884/0.2572/0.1435	0.2782/0.2363/0.1275	0.2873/0.2659/0.1408
	BSV	0.4245/0.3559/0.2282	0.3703/0.3224/0.1494	0.3270/0.2752/0.0848	0.3188/0.2809/0.0547
	DCP	0.8910/0.9410/0.8786	0.7820/0.8626/0.6980	0.5712/0.5084/0.4311	0.3051/0.2682/0.1001
	DSIMVC	–/–/–	–/–/–	–/–/–	–/–/–
Wiki_fea	**Ours**	0.5220/0.4538/0.3575	**0.4735**/0.3842/**0.2544**	**0.3967**/0.3317/0.1484	**0.3751**/0.2974/0.1172
	FreeCSL	**0.5307**/0.4471/0.3696	0.4096/0.2899/0.2042	0.3950/0.2339/**0.1650**	0.3614/0.2016/**0.1423**
	DIMVC	0.2589/0.1777/**0.3978**	0.3301/0.2183/0.1596	0.2875/0.1520/0.1000	0.2467/0.1130/0.0715
	HSACC	0.4916/0.4520/0.2657	0.4365/0.2903/0.1663	0.3657/0.2374/0.1322	0.3353/0.2629/0.0614
	DGIMVCM	0.4323/0.3556/0.2488	0.3943/0.2783/0.1596	0.3182/0.2564/0.0945	0.3745/0.2829/0.1332
	RecFormer	0.5105/**0.4757**/0.3093	0.4239/0.3833/0.1439	0.3241/0.2612/0.1122	0.3426/0.2804/0.1063
	BSV	0.5021/0.4497/0.3482	0.4522/**0.3966**/0.2377	0.3874/**0.3423**/0.1338	0.3555/**0.3194**/0.0909
	DCP	0.4477/0.4089/0.2924	0.4138/0.3080/0.2425	0.3922/0.2586/0.1320	0.3465/0.2682/0.1082
	DSIMVC	0.4581/0.3359/0.2705	0.4299/0.2877/0.2298	0.3388/0.2104/0.1409	0.3234/0.1839/0.1167
COIL20	**Ours**	**0.8021/0.8704**/0.7521	**0.7663/0.8338/0.6915**	**0.7676/0.8277/0.7086**	**0.7011/0.8029/0.6352**
	FreeCSL	0.7910/0.8586/**0.7573**	0.6944/0.8093/0.6456	0.7542/0.8248/0.6928	0.6493/0.6987/0.5378
	DIMVC	0.5826/0.7509/0.5430	0.6276/0.7253/0.5208	0.6097/0.7218/0.4886	0.3917/0.4628/0.2366
	HSACC	0.5847/0.7272/0.4334	0.6069/0.6862/0.2715	0.1639/0.2236/0.0292	0.1743/0.2208/0.0220
	DGIMVCM	0.7326/0.8645/0.6788	0.7396/0.8387/0.6884	0.7208/0.8102/0.6561	0.4826/0.6439/0.3705
	RecFormer	0.7167/0.8033/0.6536	0.7653/0.8190/0.6903	0.7319/0.7969/0.6539	0.2944/0.4398/0.1696
	BSV	0.5951/0.6914/0.4820	0.4833/0.5619/0.1965	0.3514/0.4544/0.0768	0.2882/0.3708/0.3014
	DCP	0.4910/0.7061/0.4207	0.4854/0.6938/0.3231	0.4333/0.6425/0.3052	0.2181/0.3923/0.1434
	DSIMVC	0.6139/0.7001/0.5241	0.5882/0.6772/0.4840	0.5257/0.6002/0.4021	0.4958/0.5955/0.3770

4.3 Ablation Study

We evaluate the contribution of each component by removing one module at a time while keeping the remaining pipeline fixed on two datasets across four missing rates. The full model consistently outperforms all variants, and the two QAP experts are complementary, with their performance drops peaking at different missing rates, particularly on the harder 100leaves.

Table 3: Ablation study on two datasets and the best results are highlighted in red.

		$r = 0.1$			$r = 0.3$			$r = 0.5$			$r = 0.7$		
		ACC (%)	NMI (%)	ARI (%)	ACC (%)	NMI (%)	ARI (%)	ACC (%)	NMI (%)	ARI (%)	ACC (%)	NMI (%)	ARI (%)
100leaves	**Full (Ours)**	94.23	96.16	88.70	83.31	89.58	72.91	62.54	78.24	46.01	43.81	69.26	26.87
	w/o KB-QAP	93.06	95.69	87.97	63.84	82.67	50.17	53.25	74.56	36.97	41.50	68.32	25.53
	w/o L-QAP	86.75	92.24	77.94	77.56	86.36	65.42	60.75	77.91	44.64	38.62	66.11	21.33
	w/o commonZ	84.88	91.13	76.61	75.19	84.50	61.86	61.52	76.54	45.38	42.52	68.33	26.44
HW	**Full (Ours)**	95.15	87.56	88.72	93.40	86.94	86.21	90.20	80.99	79.78	86.40	75.67	73.52
	w/o KB-QAP	95.05	87.32	86.22	92.20	84.44	83.69	89.40	79.67	78.34	86.00	75.56	72.49
	w/o L-QAP	94.15	88.81	88.41	93.08	86.46	85.47	89.78	80.48	79.19	86.15	75.40	72.67
	w/o commonZ	84.40	81.61	76.46	77.25	80.54	72.99	89.75	80.51	78.72	77.00	74.32	68.62

4.4 Convergence Analysis

To verify that the proposed framework trains stably under different missing rates, Figure 2 plots the total training loss on HW across $r \in \{0.1, 0.3, 0.5, 0.7\}$. In all four settings the loss drops sharply within the first 20 epochs and then decays smoothly toward a plateau, with no divergence or severe oscillation after epoch 80. The curves confirm that 110 training epochs are sufficient for convergence across the tested missing rates and that the dual-expert refinement does not introduce training instability.

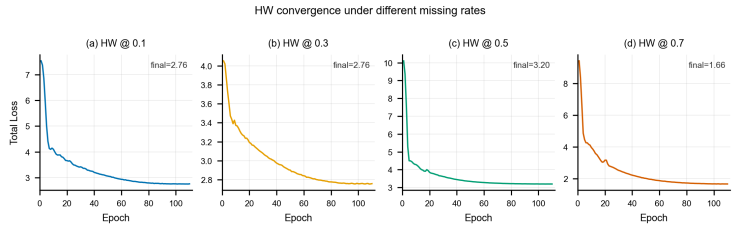

Figure 2: Training loss convergence on HW under four missing rates.

4.5 Anchor View Selection Analysis

We fix each candidate view as the anchor in turn while keeping all other components unchanged, and compare these results with the anchor selected by the learned score. Figure 3 reports ACC on HW, COIL20, and Scene15. Blue bars denote the view selected by the learned score, gray bars denote the results of fixing every other view as the anchor, and the orange dashed line denotes the average over random anchor-view choices. Across all three datasets, the automatically selected view attains the highest ACC. These results show that the learned score provides an effective criterion for selecting the anchor view.

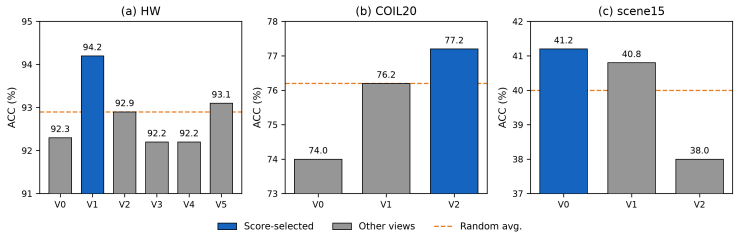

Figure 3: Anchor view selection ablation on HW, COIL20, and Scene15.

4.6 Parameter Sensitivity

We study the sensitivity of four hyperparameters directly tied to our design choices on 100leaves at $r = 0.3$, with all other settings fixed. They are the KB fusion weight w_{KB}, the number of anchors M, the KNN skeleton size k_{skel}, and the refinement decay δ. As shown in Figure 4, performance remains stable over a broad range: $w_{\mathrm{KB}} > 0.5$ consistently benefits from edge information, $M = 40$ anchors already suffice, k_{skel} is insensitive around its default, and δ peaks near 0.85 and then flattens.

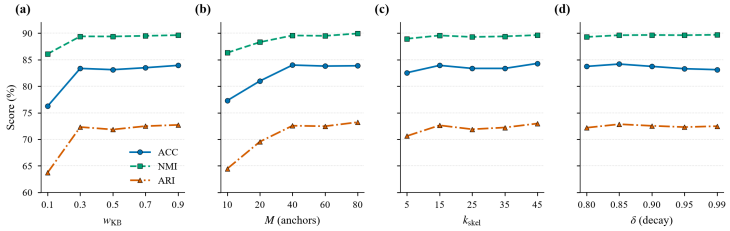

Figure 4: Parameter sensitivity on 100leaves under missing rate $r = 0.3$.

4.7 Visualization of Learned Representations

For an intuitive check of representation quality, Figure 5 visualizes the t-SNE embeddings on HW and Hdigit. Panels (a) and (c) show the raw input features, while panels (b) and (d) show our refined spectral embedding after training. The raw features are heavily mixed across classes, whereas the learned embeddings form compact and well-separated clusters with substantially higher ACC. This indicates that the anchor-guided dual-expert refinement preserves discriminative cluster structure in a shared cross-view space.

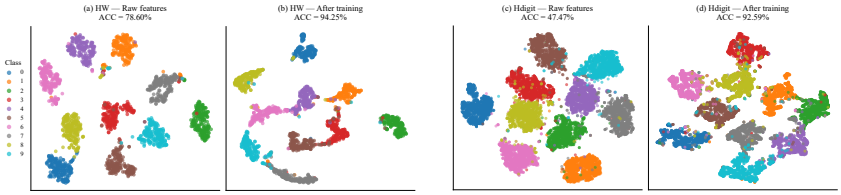

Figure 5: t-SNE visualization on HW (a, b) and Hdigit (c, d).

5 Conclusion

This paper proposes an anchored dual-expert cross-view correspondence framework for incomplete multi-view clustering. By measuring cross-view correspondence at a compact learned anchor level and aggregating the complementary evidence of two experts, the method recovers consistent cluster structure under substantial view missingness while remaining scale-friendly. Extensive experiments on seven multi-view benchmarks demonstrate clear improvements over existing imputation-free, subspace, graph, and contrastive baselines.

Limitations

Our framework has several limitations. First, we consider only the standard aligned IMC setting; extending to fully unaligned or partially view-correspondence (PVC) settings is left for future work. Second, the anchor number is dataset-dependent and manually tuned, and very large-scale datasets ($N > 50{,}000$) remain unexplored, despite anchor matching reducing the cost from $O(N^4)$ to $O(M^4)$. Third, gains diminish under extreme missingness ($r = 0.7$), as correspondence signals grow less reliable with sparse observations. Fourth, ACC improvements are consistent, but NMI and ARI gains are less uniform, with some baselines outperforming ours under specific dataset–rate conditions. Finally, evaluation is limited to image and text benchmarks; heterogeneous modalities such as graphs and point clouds are unexplored.

AI USE STATEMENT

In this work, we used large language models to assist with the following tasks: Chinese-English translation and polishing of the paper text, LaTeX formatting and layout adjustments, reference list formatting, and generation of the method overview figure (with labels and layout manually adjusted).

We did not use generative AI tools for proposing research ideas, experimental design, algorithm derivation, generation or analysis of experimental results, nor did we use AI tools to write the core technical content of the paper (including method description, experimental analysis, and conclusions). All mathematical formulas, algorithm designs, and experimental data presented in this paper were completed independently by the authors.

For all AI-assisted content, we performed manual review and verification: translated and polished text was checked sentence by sentence for technical accuracy to ensure no factual errors were introduced or original meanings altered; formatting and layout adjustments were verified through compilation; reference entries were individually verified against online databases. We take full responsibility for the final content of this work, including text and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work proposes an unsupervised learning algorithm for incomplete multi-view clustering. The research objects are publicly available multi-view benchmark datasets (including 100leaves, HW, ALOI-100, Scene15, Hdigit, Wiki, and COIL-20), which do not involve human subjects, personal privacy data, or sensitive information. All datasets used are widely used public datasets in academic research, and their collection and usage comply with relevant regulations.

The algorithm proposed in this study aims to improve the performance and scalability of multi-view clustering, has no direct potential for malicious application, and does not present significant risks related to discrimination, bias, or fairness. We confirm that this study has no conflicts of interest and has not received commercial funding that could affect the objectivity of the research. The research process adheres to academic integrity standards; all experimental results are obtained from actual runs without data manipulation or selective reporting.

REPRODUCIBILITY STATEMENT

To ensure the reproducibility of this work, we have taken the following measures:

Algorithm implementation. The anchored dual-expert cross-view correspondence framework proposed in this paper is implemented based on PyTorch 2.4.1 with Python 3.10.20 and CUDA 12.4. All network structures (including two-layer MLP view encoders, a four-head TransformerConv, KB-QAP and L-QAP expert modules) are described in detail in the method section of the paper, and the loss functions and optimization strategies are given with explicit mathematical definitions.

Experimental configuration. All hyperparameters (including learning rate, number of training epochs, number of anchors, Sinkhorn iterations, number of nearest neighbors, etc.) are described in the implementation details section and tuned for each dataset. Missing patterns are generated with five different random seeds, and all methods run under the same missing masks; for each mask, every method is run five times with different random initializations using the same initialization seeds across methods. We report the mean of ACC, NMI, and ARI.

Datasets. The 7 multi-view benchmark datasets used in this paper are all publicly available datasets.

Code release. We plan to release the complete source code, training scripts, and configuration files after the paper is accepted, to facilitate reproduction and comparison by researchers.

REFERENCES

Yuzhuo Dai, Jiaqi Jin, Zhibin Dong, Siwei Wang, Xinwang Liu, En Zhu, Xihong Yang, Xinbiao Gan, and Yu Feng. Imputation-free and alignment-free: Incomplete multi-view clustering driven by consensus semantic learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5071–5081, 2025.

Xiaojian Ding, Lin Zhao, Xian Li, and Xiaoying Zhu. Incomplete multi-view clustering via hierarchical semantic alignment and cooperative completion. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, pp. 58096–58116, 2025.

Yiming Du, Ziyu Wang, Jian Li, Rui Ning, and Lusi Li. Deep incomplete multi-view clustering via hierarchical imputation and alignment. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 40, pp. 20941–20949, 2026.

Uno Fang, Man Li, Jianxin Li, Longxiang Gao, Tao Jia, and Yanchun Zhang. A comprehensive survey on multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 35 (12):12350–12368, 2023.

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)*, volume 37, 2024.

Yaoyuan Guo, Zhibin Gu, Songhe Feng, Yuhui Zheng, and Bing Li. Sport: Structure-aware prototype disentanglement for incomplete multi-view clustering. *arXiv preprint arXiv:2607.10413*, 2026.

Hongqing He, Jie Xu, Wenyuan Yang, Yonghua Zhu, Guoqiu Wen, and Xiaofeng Zhu. Global-graph guided and local-graph weighted contrastive learning for unified clustering on incomplete and noise multi-view data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026.

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 (CVPR)*, pp. 11600–11609, 2023.

Yijie Lin, Yuanbiao Gou, Zitao Liu, Boyun Li, Jiancheng Lv, and Xi Peng. Completer: Incomplete multi-view clustering via contrastive prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11169–11178, 2021.

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, 2023.

Yijie Lin, Mouxing Yang, Peng Hu, Jiancheng Lv, Hao Chen, and Xi Peng. Learning with partial and noisy correspondence in graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 48(7):8234–8249, 2026. doi: 10.1109/TPAMI.2026.3670236.

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

Huayi Tang and Yong Liu. Deep safe incomplete multi-view clustering: Theorem and algorithm. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2022.

Yiming Wang, Qun Li, Dongxia Chang, Jie Wen, Hua Dai, Fu Xiao, and Yao Zhao. Learning from disjoint views: A contrastive prototype matching network for fully incomplete multi-view clustering. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, 2025.

Jie Wen, Ke Yan, Zheng Zhang, Yong Xu, Junqian Wang, and Lunke Fei. Adaptive graph completion based incomplete multi-view clustering. *IEEE Transactions on Multimedia*, 23:2493–2504, 2021.

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 *International Conference on Machine Learning (ICML)*, 2024.

Yanwanyu Xi, Xiao Zheng, Chang Tang, Xingchen Hu, Yuanyuan Liu, Jun-Jie Huang, and Xinwang Liu. Lrgr: Self-supervised incomplete multi-view clustering via local refinement and global realignment. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2025.

Gehui Xu, Jie Wen, Chengliang Liu, Bing Hu, Yicheng Liu, Lunke Fei, and Wei Wang. Deep variational incomplete multi-view clustering: Exploring shared clustering structures. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 16147–16155, 2024.

Jie Xu, Chao Li, Yazhou Ren, Liang Peng, Yujie Mo, Xiaoshuang Shi, and Xiaofeng Zhu. Deep incomplete multi-view clustering via mining cluster complementarity. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2022.

Jun Yin, Pei Wang, Shiliang Sun, and Zhonglong Zheng. Incomplete multi-view clustering via multi-level contrastive learning. *IEEE Transactions on Knowledge and Data Engineering*, 37(8): 4716–4727, 2025.

Shengju Yu, Zhibin Dong, Siwei Wang, Pei Zhang, Yi Zhang, Xinwang Liu, Naiyang Guan, Tiejun Li, and Yiu-ming Cheung. Simple yet effective incomplete multi-view clustering: Similarity-level imputation and intra-view hybrid-group prototype construction. In *International Conference on Learning Representations (ICLR)*, 2025.

Pengxin Zeng, Mouxing Yang, Yiding Lu, Changqing Zhang, Peng Hu, and Xi Peng. Semantic invariant multi-view clustering with fully incomplete information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(4):2139–2150, 2024.

Zhenghao Zhang, Jun Xie, Xingchen Chen, Tao Yu, Hongzhu Yi, Kaixin Xu, Yuanxiang Wang, Tianyu Zong, Xinming Wang, Jiahuan Chen, Guoqing Chao, Feng Chen, Zhepeng Wang, and Jungang Xu. Dynamic deep graph learning for incomplete multi-view clustering with masked graph reconstruction loss. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2026.

Shuping Zhao, Jie Wen, Lunke Fei, and Bob Zhang. Tensorized incomplete multi-view clustering with intrinsic graph completion. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 37, 2023.

Table 4: Main global hyperparameters fixed across all datasets.

Parameter	Value	Role
α, β, ζ	0.6, 0.2, 0.2	Anchor view scoring weights
λ_{edge} (KB)	0.3	Node/edge affinity mixing
$\lambda_{\text{within}}, \lambda_{\text{sym}}$ (KB)	0.1, 0.1	Structure consistency / symmetry
κ (KB)	5.0	BCE logit temperature
λ_L	1.0	Matching-term weight in total loss
$\omega_0 \to \omega_1$	$1.0 \to 0.1$	L-QAP edge-affinity annealing
τ_{SK}	0.1	Sinkhorn temperature
$\lambda_{\text{edge}}, \lambda_{\text{sym}}$ (L-QAP)	0.1, 0.1	Edge consistency / symmetry
w_{out}, b	5.0, -10.0	Outlier BCE weight / background logit
μ	$\min(0.95,\ 0.85 + 0.15\bar{r})$	Pseudo-label momentum
w_p	0.5	Partially-consistent update weight
w_{KB}, w_{LQAP}	0.5, 0.5	Dual-expert affinity fusion

Table 5: Dataset-specific hyperparameters. C is the number of clusters, h the encoder hidden widths, d the embedding dimension, and r the view-wise missing rate. k_{sel} is the KNN size for anchor-view selection, τ_v the initial pseudo-label similarity quantile, τ_{init} the training-threshold initial value, τ_c the contrastive temperature, r_c the positive-set fraction, k_{edge} the edge-affinity KNN, M the number of learnable anchors, k_{skel} the L-QAP skeleton KNN, k_{clu} the spectral-clustering KNN, and η the learning rate. All unlisted hyperparameters are shared across datasets.

Dataset	C	h	d	r	Anchor & pseudo-label			Contrastive		QAP graph			k_{clu}	η
					k_{sel}	τ_v	τ_{init}	τ_c	r_c	k_{edge}	M	k_{skel}		
100leaves	100	$[512, 256]$	512	0.1	20	0.95	0.99	0.80	0.10	15	40	25	8	10^{-4}
				0.3	20	0.95	0.9	1.00	0.10	8	40	25	6	10^{-4}
				0.5	20	0.95	0.99	0.80	0.10	8	40	25	4	10^{-4}
				0.7	20	0.90	0.99	0.60	0.10	8	40	25	4	10^{-4}
HW	10	$[256, 128]$	64	0.1	35	0.95	0.95	0.80	0.10	5	40	10	30	10^{-4}
				0.3	35	0.90	0.95	0.80	0.10	5	40	10	30	10^{-4}
				0.5	35	0.85	0.95	0.80	0.05	5	40	10	30	4×10^{-4}
				0.7	35	0.70	0.95	0.15	0.06	5	40	10	10	4×10^{-4}
ALOI-100	100	$[512, 256]$	512	0.1	50	0.99	0.99	0.80	0.10	60	80	40	80	10^{-4}
				0.3	50	0.85	0.95	0.80	0.10	20	100	40	30	10^{-4}
				0.5	40	0.80	0.95	0.50	0.08	20	100	25	20	10^{-4}
				0.7	40	0.75	0.99	0.30	0.05	20	100	25	30	10^{-4}
Scene15	15	$[128, 64]$	128	0.1	15	0.99	0.95	0.80	0.15	10	60	40	35	8×10^{-5}
				0.3	15	0.95	0.95	0.50	0.15	15	60	40	35	8×10^{-5}
				0.5	15	0.95	0.95	0.50	0.15	15	60	40	35	8×10^{-5}
				0.7	15	0.95	0.95	0.50	0.15	15	60	40	25	2.5×10^{-4}
Hdigit	10	$[512, 256]$	256	0.1	50	0.95	0.99	0.80	0.10	60	80	40	25	2×10^{-4}
				0.3	50	0.90	0.95	0.60	0.10	60	80	40	25	2×10^{-4}
				0.5	50	0.80	0.95	0.40	0.10	60	80	40	25	10^{-4}
				0.7	50	0.75	0.85	0.20	0.15	60	80	40	25	10^{-4}
Wiki_fea	10	$[256, 128]$	128	0.1	30	0.98	0.98	0.80	0.10	25	50	20	25	10^{-4}
				0.3	30	0.90	0.95	0.60	0.10	25	50	20	20	10^{-4}
				0.5	30	0.85	0.95	0.50	0.10	25	50	20	20	2.5×10^{-4}
				0.7	30	0.70	0.90	0.20	0.10	20	50	20	15	10^{-4}
COIL20	20	$[256, 128]$	128	0.1	20	0.95	0.99	0.80	0.10	20	40	20	15	1.5×10^{-4}
				0.3	20	0.90	0.98	0.60	0.10	20	40	20	15	10^{-4}
				0.5	20	0.85	0.98	0.50	0.10	20	40	20	15	10^{-4}
				0.7	20	0.80	0.98	0.20	0.10	20	40	20	15	10^{-4}