HIGH-ORDER CONSENSUS COMPLETION FOR INCOMPLETE CONTRASTIVE MULTI-VIEW CLUSTERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Incomplete contrastive multi-view clustering aims to learn discriminative representations from incomplete multi-view data. However, existing methods still have two issues: 1) Most imputation-based methods mainly rely on local cross-view relations for missing-view recovery, while neglecting high-order relationships among samples, which may lead to unreliable completion. 2) View incompleteness obscures the latent relationships among samples, making it difficult to expand reliable positive pairs beyond same-sample cross-view pairs. To address these issues, we propose a framework named high-order consensus completion for incomplete contrastive multi-view clustering (HCCIMC). Specifically, we design a high-order consensus completion module, which first constructs a global consensus graph from view-specific bipartite graphs. Then, multi-order random walks are performed on the consensus graph to derive high-order neighborhood structures, which are adaptively combined to recover missing-view representations. Furthermore, we propose a multi-order neighborhood consistency contrastive learning module. It models multi-order neighborhood relationships and evaluates their consistency across different views to select reliable neighborhood samples as positives. These two modules mutually enhance each other and help learn robust clustering-friendly representations. Extensive experimental results demonstrate that our method achieves the state-of-the-art clustering performance.

1 INTRODUCTION

Multi-view data refer to multiple representations of the same objects obtained from different sensors or feature extractors (Fang et al., 2023; Li et al., 2019). Compared with single-view data, multi-view data integrate complementary information from different views, enabling a more comprehensive characterization of the underlying data structure. Multi-view clustering aims to mine the consistent and complementary information across views to improve clustering performance (Wu et al., 2024b; Xu et al., 2022b). However, during data acquisition, some samples inevitably suffer from incomplete views due to sensor instability or other external disturbances. Since missing data are common in practical applications, incomplete multi-view clustering (IMVC) has attracted increasing attention in recent years (Du et al., 2026; Dong et al., 2025b). The mainstream deep incomplete multi-view clustering (DIMVC) methods can be broadly divided into two categories: imputation-free methods and imputation-based methods (Zhang et al., 2025).

Imputation-free methods typically learn latent representations solely from available views without explicitly recovering missing views. For example, APADC proposed an adaptive feature projection mechanism with distribution alignment, mapping view-specific features into a common space to reduce the impact of missing data (Xu et al., 2023). Dai et al. (2025) proposed FreeCSL, an imputation- and alignment-free IMVC framework that learns consensus prototypes from available data to construct a shared semantic space . However, under high missing rates, imputation-free methods often suffer from degraded performance because limited available views provide insufficient cross-view complementary and consensus information.

In contrast, imputation-based methods recover missing data or latent representations by exploiting cross-view dependencies, and then perform clustering (Lin et al., 2021; Liu et al., 2023; Wang et al., 2021a). CRTC recovered missing data by transferring inter-instance similarity relationships from observed views to missing views (Wang et al., 2023). BURG introduced a flow-based dis-

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.

tribution transfer mechanism that integrates observed-view distributions and transfers the learned consensus distribution to missing views (Jin et al., 2025). However, these methods usually recover missing data primarily through local cross-view relations, while insufficiently exploiting high-order neighborhood information, which may lead to unreliable recovery (Li et al., 2022; Xu et al., 2025). In addition, contrastive learning is widely used in DIMVC for unified representation learning (Zhuang et al., 2025). However, conventional multi-view contrastive learning treats representations of the same sample across views as positives, while view incompleteness reduces such pairs. Although neighboring samples can expand the positive set, incomplete views may yield unreliable neighborhood relations, making reliable positive selection challenging.

To address the above issues, we exploit high-order neighborhood information among samples, which captures more reliable inter-sample relationships by aggregating structural information along multi-step paths. To provide an intuitive illustration, we construct KNN graphs on the TwoMoon dataset under different propagation orders, as shown in Figure 1. As the propagation order increases, the number of wrong cross-class connections (red edges) gradually decreases from 102 to 7, while the node accuracy increases from 0.86 to 0.97. It is evident that high-order neighborhood information effectively reduces erroneous cross-class relations.

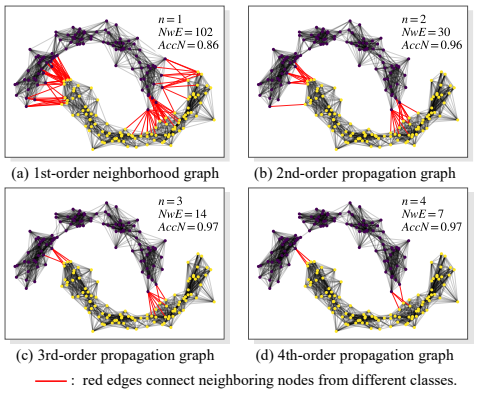


Figure 1: KNN graphs on the TwoMoon dataset under different propagation orders. Red edges indicate wrong cross-class connections. NwE denotes the number of wrong edges, and AccN denotes the node accuracy.

Motivated by these observations, we propose a framework named High-Order Consensus Completion for Incomplete Contrastive Multi-view Clustering (HCCIMC), which exploits high-order neighborhood information for missing-view completion and reliable positive pairs construction. Specifically, HCCIMC first constructs view-specific bipartite graphs from available latent representations and integrates them into a global consensus graph. Based on this graph, multi-order random walks are performed to capture high-order neighborhood structures, which are adaptively aggregated to complete missing-view representations. Furthermore, HCCIMC introduces a multi-order neighborhood consistency contrastive learning strategy. This strategy models multi-order neighborhood relationships and evaluates their consistency across different views. Reliable neighborhood samples are then selected as positives to reduce false positive pairs. In brief, the major contributions are summarized as follows:

- We propose a high-order consensus completion mechanism, which constructs a global consensus graph from view-specific bipartite structures and adaptively aggregates high-order neighborhood structures for reliable missing-view representation completion.
- We propose a contrastive learning strategy based on multi-order neighborhood consistency, which integrates first- and high-order neighborhood relations and verifies their cross-view consistency to select reliable positive pairs.
- Extensive experiments are conducted on four public datasets and the results demonstrate the state-of-the-art clustering performance of our proposed method.

2 Related Work

2.1 Deep Incomplete Multi-view Clustering

DIMVC methods can be broadly divided into two categories: imputation-free methods and imputation-based methods. For imputation-free IMVC methods, they usually perform clustering directly on the available data (Feng et al., 2024; Wang et al., 2021b). For instance, Xu et al. (2022a) map complete data embeddings into a high-dimensional space to mine cluster complementarity and transform the complementary information into high-confidence supervision for clustering

consistency. Xue et al. (2025) learn robust unified representations from incomplete multi-view data through adaptive view fusion and multi-graph contrastive regularization . However, imputation-free methods mainly rely on available views, which limits their clustering performance. Imputation-based IMVC methods usually attempt to recover missing data or latent features first, and then perform clustering on the completed multi-view data (Li et al., 2023). Huang et al. (2024) generate simulated candidates based on view-specific statistics and employ a data evaluator to select appropriate candidates for missing-view recovery . Chao et al. (2024) transfer cross-view graph relations and employ GCN-based message passing to recover missing data . Similarly, Lu et al. (2024) use cross-view decoders to recover missing-view representations . Most imputation-based methods rely on local cross-view relations to recover missing data, with insufficient consideration of high-order neighborhood information, which may lead to unreliable recovered information.

2.2 Contrastive Learning

In recent years, contrastive learning has become a popular unsupervised representation learning paradigm (Chen et al., 2020; Xiao et al., 2025). Its core idea is to learn discriminative representations by pulling positive pairs and pushing negative pairs, which has greatly contributed to the development of incomplete multi-view clustering. For instance, DCP employs contrastive learning to maximize cross-view mutual information for consistent representation learning and adopts dual prediction to recover missing views (Lin et al., 2023). RGCL introduces noise-robust graph contrastive learning, treating in-graph samples as absolute positives, highly similar out-of-graph samples as semi-positives, and the rest as negatives (Zhuang et al., 2025). RMGC introduces a structured RINCE mechanism that assigns similarity-based weights to negative pairs, reducing the influence of false negatives (Xue et al., 2025). Although contrastive learning has enhanced representation learning in IMVC, reliable positive pairs identification remains challenging because missing views disrupt cross-view correspondences and neighborhood relationships.

3 The Proposed Method

3.1 Problem Definition

In this paper, we define a multi-view dataset as $\mathcal{X} = \{\mathbf{X}^v\}_{v=1}^{V}$, where V is the number of views. The feature matrix of the v-th view is denoted by $\mathbf{X}^v = \{x_i^v\}_{i=1}^{N} \in \mathbb{R}^{N \times d_v}$, where N is the number of samples and d_v is the feature dimension of the v-th view. To characterize the incompleteness of multi-view data, we introduce an indicator matrix $\mathbf{M} \in \{0,1\}^{N \times V}$, where $M_{iv} = 1$ means the i-th sample exists in the v-th view, otherwise $M_{iv} = 0$. Figure 2 illustrates the overall framework of HCCIMC.

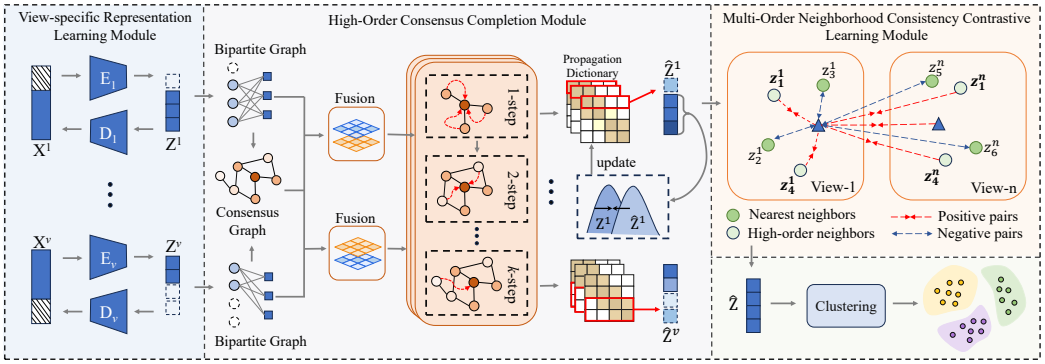


Figure 2: The overall framework of the proposed HCCIMC.

3.2 VIEW-SPECIFIC REPRESENTATION LEARNING MODULE

For multi-view clustering, learning consensus and complementary features across views is essential. Many studies have demonstrated the superior capability of autoencoders in feature extraction (Vincent et al., 2010). Accordingly, we introduce several view-specific autoencoders. Each encoder extracts latent features from its corresponding view, while the corresponding decoder reconstructs the input. Specifically, for the v-th view, let E_θ^v represent the corresponding encoder. The latent representation of the i-th sample extracted by the encoder of the v-th view can be formulated as:

$$z_i^v = E_\theta^v(x_i^v), \tag{1}$$

where z_i^v denotes the latent representation of the i-th sample in the v-th view. We collect the latent representations of all samples into $\mathbf{Z}^v \in \mathbb{R}^{N\times d}$, where the rows corresponding to missing samples are set to zero. Similarly, for the v-th view, the decoder process of the i-th sample can be formulated as:

$$\hat{x}_i^v = D_\phi^v(z_i^v), \tag{2}$$

where $\hat{x}_i^v$ denotes the reconstructed data of the i-th sample in the v-th view. To train the autoencoders for all views, the reconstruction loss over the samples can be formulated as:

$$\mathcal{L}_{rec} = \sum_{v=1}^{V}\sum_{i=1}^{N} M_{iv} \left\| x_i^v - \hat{x}_i^v \right\|_2^2 . \tag{3}$$

The indicator M_{iv} ensures that only the available samples contribute to the reconstruction loss.

3.3 HIGH-ORDER CONSENSUS COMPLETION MODULE

To recover missing-view representations more reliably, we propose a high-order consensus completion module (HCCM). HCCM integrates view-specific structural graphs into a global consensus graph and performs multi-order random walks to capture high-order neighborhood information. The resulting multi-order representations are adaptively aggregated for missing-view completion.

To construct the view-specific structural graph, we introduce an anchor-based bipartite graph to compactly characterize sample relationships (Guo & Ye, 2019). Specifically, k-means is performed on the available samples in each view to obtain K_a representative anchors. These anchors act as local structural prototypes that summarize the intra-view relationships among samples. Subsequently, the bipartite graph of the v-th view is initialized as $\mathbf{G}^v \in \mathbb{R}^{N\times K_a}$. The affinity matrix of the v-th view can be formulated as:

$$\mathbf{S}^v = (\mathbf{D}^v)^{-\frac{1}{2}} \mathbf{G}^v \mathbf{G}^{v\top} (\mathbf{D}^v)^{-\frac{1}{2}} , \tag{4}$$

where $\mathbf{D}^v$ is the diagonal degree matrix induced by $\mathbf{G}^v$. To suppress noisy weak connections, we further sparsify $\mathbf{S}^v$ by retaining the top-k largest affinities for each sample. Then, the affinity matrices of all views are aggregated into a global affinity matrix as follows:

$$S_{ij}^{\text{cons}} = \frac{\sum_{v=1}^{V} M_{iv} M_{jv} S_{ij}^v}{\sum_{v=1}^{V} M_{iv} M_{jv} + \epsilon}, \tag{5}$$

where ϵ is used to avoid division by zero.

High-order neighborhood relationships can better characterize complex sample relationships (Pan & Kang, 2023). We first fuse the global consensus affinity matrix with the view-specific affinity matrix and row-normalize the result to obtain the transition matrix:

$$\hat{S}_{ij}^v = \frac{S_{ij}^{\text{cons}} + S_{ij}^v}{\mathbb{I}[S_{ij}^{\text{cons}} > 0] + \mathbb{I}[S_{ij}^v > 0]}, \tag{6}$$

$$P_{ij}^v = \frac{\hat{S}_{ij}^v}{\sum_{k=1}^{N} \hat{S}_{ik}^v}, \tag{7}$$

where $\mathbb{I}[\cdot]$ is the indicator function, which equals 1 if the condition holds and 0 otherwise. Based on the transition matrix $\mathbf{P}^v$, high-order neighborhood information is progressively propagated through

multi-step random walks. We initialize $\mathbf{H}_0^v = \mathbf{Z}^v$. The propagated representation after the t-th random-walk step is then computed as:

$$\mathbf{H}_t^v = \mathbf{P}^v \mathbf{H}_{t-1}^v, \quad t = 1, 2, \ldots, T, \tag{8}$$

$$\mathcal{T}^v = \left[\mathbf{P}^v \mathbf{H}_0^v,\ (\mathbf{P}^v)^2 \mathbf{H}_0^v,\ \ldots,\ (\mathbf{P}^v)^T \mathbf{H}_0^v\right], \tag{9}$$

where $\mathbf{H}_t^v$ denotes the representation after t random-walk steps, $\mathcal{T}^v$ collects the propagated representations from different orders, and T is the maximum propagation order.

To adaptively integrate structural information from different propagation orders, we introduce an observation-constrained calibration strategy. The key idea is that, on the observed sample set Ω_v of the v-th view, the weighted combination of multi-order propagated representations should be close to the original latent representations. The learned weights are then used to aggregate multi-order propagated information for missing-view representation completion. The calibration objective is formulated as:

$$\begin{aligned} \arg\min_{\mathbf{w}^v} \quad & \left\| \mathcal{T}_{\Omega_v}^v \mathbf{w}^v - \mathbf{H}_{0,\Omega_v}^v \right\|_F^2 + \lambda \|\mathbf{w}^v\|_2^2 \\ & + \gamma \sum_{t=1}^{T} w_t^v \log w_t^v \\ \text{s.t.} \quad & \sum_{t=1}^{T} w_t^v = 1, \quad w_t^v \geq \delta, \end{aligned} \tag{10}$$

where $\mathcal{T}_{\Omega_v}^v$ denotes the multi-order propagation dictionary on the available sample set Ω_v, and $\mathbf{H}_{0,\Omega_v}^v$ denotes the corresponding original latent representations. Additionally, λ and γ are regularization coefficients. δ denotes the lower bound of the weights.

The missing latent representations in the v-th view are completed as follows:

$$\hat{\mathbf{Z}}^v = \mathbf{M}_{\cdot v} \odot \mathbf{Z}^v + (\mathbf{1} - \mathbf{M}_{\cdot v}) \odot \left(\sum_{t=1}^{T} w_t^v \mathbf{H}_t^v \right). \tag{11}$$

where $\mathbf{M}_{\cdot v}$ denotes the v-th column of the observation indicator matrix, and $\odot$ denotes the element-wise product with broadcasting along the feature dimension.

3.4 Multi-Order Neighborhood Consistency Contrastive Learning

Reliable positive-pair construction is crucial for contrastive learning (Ding et al., 2025). We propose a multi-order neighborhood consistency contrastive learning (MNCCL) module to construct reliable positive pairs. Samples that lie close to each other in the feature space are more likely to belong to the same category. Therefore, we introduce neighborhood information to guide positive and negative pair selection. For the v-th view, the construction of the k-nearest neighbor graph can be formulated as:

$$A_{ij}^v = \begin{cases} 1, & \left(\hat{z}_i^v \in \Lambda(\hat{z}_j^v)\right) \text{ and } \left(\hat{z}_j^v \in \Lambda(\hat{z}_i^v)\right), \\ 0, & \text{otherwise}, \end{cases} \tag{12}$$

where $\Lambda(\hat{z}_i^v)$ denotes the neighbor set of $\hat{z}_i^v$ in the v-th view. Subsequently, to capture more robust structural information, we introduce high-order neighborhood relations. Inspired by the good-neighbor strategy in (Wu et al., 2024a), we measure the Jaccard similarity between first-order neighbor sets, which is defined as follows:

$$\mathrm{Jac}^v(i,j) = \frac{\left|\Lambda(\hat{z}_i^v) \cap \Lambda(\hat{z}_j^v)\right|}{\left|\Lambda(\hat{z}_i^v) \cup \Lambda(\hat{z}_j^v)\right|}, \tag{13}$$

If two samples share sufficient neighbors, they are regarded as having a higher-order neighborhood relation. Accordingly, the high-order neighborhood graph in the v-th view can be formulated as:

$$B_{ij}^v = \begin{cases} 1, & \mathrm{Jac}^v(i,j) > f, \\ 0, & \text{otherwise}, \end{cases} \tag{14}$$

where f is a predefined threshold for selecting higher-order neighbors. The multi-order neighborhood graph in the v-th view is formulated as:

$$R_{ij}^{v} = A_{ij}^{v} \vee B_{ij}^{v}, \tag{15}$$

where $\vee$ denotes the logical OR operation.

Although R^v provides multi-order neighborhood positives within each view, different views have distinct neighborhood structures. To select more reliable positives, we require semantic neighbors to exhibit consistent neighborhood relations across views. This cross-view consistency is formulated as:

$$C_{ij}^{vu} = \begin{cases} 1, & R_{ij}^{v} = 1 \text{ and } R_{ij}^{u} = 1, \\ 0, & \text{otherwise}, \end{cases} \tag{16}$$

where C_{ij}^{vu} indicates whether samples i and j are consistently identified as neighbors in both views v and u. Based on C^{vu} , a reliable positive sample set $\mathcal{P}_i^{vu}$ can be constructed as:

$$\mathcal{P}_i^{vu} = \left\{ j \mid j = i \vee \ C_{ij}^{vu} = 1 \right\}, \tag{17}$$

where $j = i$ denotes the cross-view correspondence of the same instance. The inter-view contrastive loss is defined as:

$$\mathcal{L}^{vu} = -\frac{1}{N} \sum_{i=1}^{N} \log \frac{\sum_{j \in \mathcal{P}_i^{vu}} \exp\left(s(\hat{z}_i^{v}, \hat{z}_j^{u})/\tau_c\right)}{\sum_{j=1}^{N} \exp\left(s(\hat{z}_i^{v}, \hat{z}_j^{u})/\tau_c\right)}, \tag{18}$$

where τ_c is the temperature parameter and $s(z_i, z_j) = z_i^{\top} z_j / (\|z_i\| \|z_j\|)$ denotes the cosine similarity. The overall contrastive learning loss can be formulated as:

$$\mathcal{L}_{cl} = \frac{1}{2} \sum_{v=1}^{V} \sum_{u \neq v} \mathcal{L}^{vu}. \tag{19}$$

3.5 The Overall Loss Function

The overall loss function for the entire network can be expressed as follows:

$$\mathcal{L} = \alpha \mathcal{L}_{rec} + \beta \mathcal{L}_{cl}. \tag{20}$$

where α and β are two hyper-parameters that control the balance between the two loss components. The complete learning algorithm and computational complexity analysis are provided in the supplementary material.

4 Experiments

4.1 Experimental Settings

We perform our proposed HCCIMC using PyTorch 2.2.2 and conduct all experiments on Windows with an NVIDIA 4060 GPU and 24GB RAM. For all datasets, we use fully connected networks to achieve autoencoders for all views with the layer dimensions of d_v-500-1000-2000-128 (128-2000-1000-500-d_v) as the encoder (decoder). Moreover, ReLU serves as the activation function, and Adam is selected as the optimizer with a default learning rate of 0.0003. The batch size is 256. We train the proposed HCCIMC network in two stages: warm-up and fine-tuning. In the warm-up stage, all view-specific autoencoders are trained for 200 epochs. In the fine-tuning stage, we set the number of anchors K_a to 50, the Jaccard similarity threshold f to 0.8, and the temperature parameter τ_c to 0.1. For each view, the number of neighbors K_n used to construct the nearest neighbor graph is selected from $\{3, 5, 7, 10\}$. The number of random-walk steps is set to 4. Additionally, the values of the two hyper-parameters α and β are chosen from the set $\{0.0001, 0.001, 0.01, 0.1, 1.0\}$. More implementation details are provided in the supplementary material.

4.2 Datasets and Evaluation Metrics

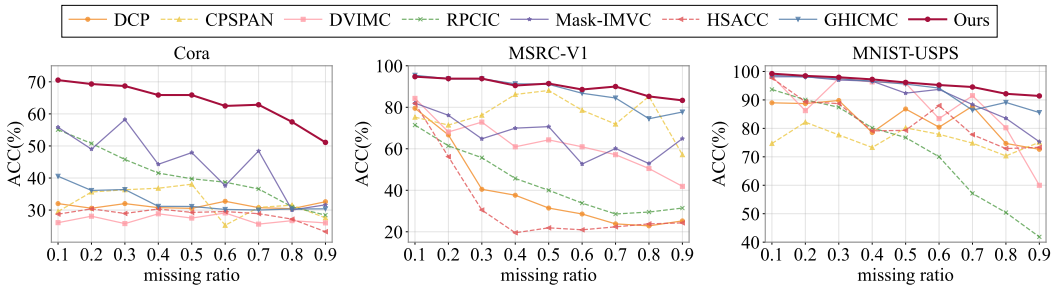


Figure 3: Comparison of the clustering performance on three datasets as the missing rate increases.

To evaluate the effectiveness of the proposed method, we conduct extensive experiments on four widely used benchmark datasets: Cora (Fang et al., 2024), Hdigit (Chen et al., 2022), MNIST-USPS (Peng et al., 2019), and MSRC-V1 (Winn & Jojic, 2005). For the incomplete multi-view setting, the incomplete rate is defined as $\eta = (N - N_c)/N$, where N_c is the number of complete samples. For each incomplete sample, we randomly delete its views. Please see the supplementary material for the detailed construction of incomplete multi-view data. Table 1 provides detailed statistics of these datasets. Three metrics, ACC, NMI, and ARI, are adopted to evaluate the clustering performance. We compare HCCIMC with six imputation-based methods, DCP (Lin et al., 2023), CPSPAN (Jin et al., 2023), RPCIC (Yuan et al., 2024), SCVT (Dong et al., 2025a), GHICMC (Chao et al., 2025), and HSACC (Ding et al., 2026), and two imputation-free methods, DVIMC (Xu et al., 2024) and Mask-IMvC (Li et al., 2025), under their recommended settings.

Table 1: Description of benchmark datasets.

Dataset	Samples	Clusters	Views	Dimensions
Cora	2708	7	4	2708/1433/2708/2708
Hdigit	10000	10	2	784/256
MNIST-USPS	5000	10	2	784/784
MSRC-V1	210	7	6	1302/48/512/100/256/210

4.3 Experimental Results and Analysis

Table 2 reports the clustering results of different IMVC methods. Overall, HCCIMC achieves superior performance on the four datasets. For example, on Cora with a missing rate of 0.7, HCCIMC outperforms the second-best method by 8.75%, 6.2%, and 12.45% in ACC, NMI, and ARI, respectively. Compared with imputation-free methods, HCCIMC consistently achieves better performance, indicating that recovering latent missing information is crucial for learning robust clustering representations. Compared with imputation-based methods, HCCIMC achieves more stable performance by exploiting high-order consensus structures, which demonstrates the superiority of the proposed high-order completion strategy. Moreover, the multi-order neighborhood consistency strategy constructs more reliable positive pairs, further improving clustering performance. Figure 3 shows the ACC curves of different methods as the missing rate increases from 0.1 to 0.9. As the missing rate increases, the ACC of most methods decreases to varying degrees. HCCIMC also exhibits some degradation under high missing rates, but it consistently achieves better overall performance than most baselines, demonstrating its robustness in incomplete multi-view scenarios.

4.4 Parameter Sensitivity Analysis

We conduct sensitivity analysis of hyper-parameters α, β, the random-walk steps T, and the number of nearest neighbor samples K_n on Cora, Hdigit, and MSRC-V1. When performing sensitivity experiments for a single parameter, all other parameters are fixed at their optimal values. As shown in Figure 4, HCCIMC shows relatively stable clustering performance under different combinations of α and β. The number of random-walk steps T controls the range of high-order neighborhood propagation. As shown in Figure 5(a), the clustering performance remains relatively stable when T varies from 1 to 6, particularly under low missing rates. Although larger fluctuations are observed under high missing rates, the performance does not consistently decrease as T increases. This indicates that the observation-constrained calibration strategy can adaptively weight different propagation orders and suppress unreliable high-order information. We uniformly set T to 4 on all datasets. As

Table 2: Clustering performance comparison on four incomplete multi-view datasets. The best and second-best results are highlighted in bold and underline, respectively.

	Method	0.1			0.3			0.5			0.7		
		ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
Cora	DCP	0.3202	0.0945	0.0389	0.3202	0.0783	0.0360	0.3054	0.0129	0.0011	0.3076	0.0191	0.0023
	CPSPAN	0.2954	0.1575	0.0754	0.3634	0.1716	0.1169	0.3807	0.1549	0.1062	0.3083	0.1149	0.0751
	DVIMC	0.2607	0.0104	0.0137	0.2578	0.0128	0.0021	0.2751	0.0107	0.0078	0.2563	0.0101	0.0102
	RPCIC	0.5506	0.4067	0.2922	0.4579	0.2939	0.1831	0.3977	0.2287	0.1143	0.3660	0.1631	0.0638
	Mask-IMvC	0.5582	0.3646	0.2589	0.5823	0.3803	0.3401	0.4792	0.2973	0.1980	0.4844	0.3104	0.2076
	HSACC	0.2873	0.0052	0.0059	0.2895	0.0063	0.0053	0.2928	0.0085	0.0098	0.2888	0.0054	0.0067
	SCVT	0.6521	0.3954	0.3908	0.6167	0.3591	0.3364	0.6134	0.3448	0.3280	0.5410	0.2617	0.2364
	GHICMC	0.4055	0.1424	0.1606	0.3641	0.0799	0.0360	0.3113	0.0461	0.0071	0.3002	0.0098	0.0066
	HCCIMC	**0.7053**	**0.5181**	**0.5001**	**0.6869**	**0.4735**	**0.4472**	**0.6588**	**0.4254**	**0.3991**	**0.6285**	**0.3724**	**0.3609**
Hdigit	DCP	0.9501	0.8843	0.8927	0.9427	0.8649	0.8772	0.9232	0.8242	0.8371	0.8851	0.7587	0.7609
	CPSPAN	0.9663	0.9169	0.9268	0.9569	0.8983	0.9072	0.9572	0.8976	0.9076	0.9458	0.8737	0.8841
	DVIMC	0.9904	0.9726	0.9788	0.9807	0.9489	0.9579	0.9681	0.9176	0.9305	0.9634	0.9074	0.9207
	RPCIC	0.9658	0.9146	0.9257	0.9380	0.8640	0.8675	0.9345	0.8545	0.8605	0.6288	0.5844	0.4510
	Mask-IMvC	0.9887	0.9673	0.9751	0.9663	0.9160	0.9262	0.9304	0.8584	0.8480	0.9056	0.8132	0.7988
	HSACC	0.9435	0.8821	0.8799	0.9477	0.8840	0.8882	0.9443	0.8687	0.8807	0.9230	0.8233	0.8374
	SCVT	0.9907	0.9723	0.9794	0.9728	0.9303	0.9405	0.7857	0.6978	0.5509	0.6397	0.6115	0.4167
	GHICMC	0.9899	0.9698	0.9777	0.9769	0.9364	0.9495	0.9641	0.9067	0.9221	0.9461	0.8700	0.8846
	HCCIMC	**0.9923**	**0.9763**	**0.9830**	**0.9847**	**0.9560**	**0.9663**	**0.9745**	**0.9308**	**0.9442**	**0.9641**	**0.9083**	**0.9221**
MNIST-USPS	DCP	0.8898	0.9245	0.8712	0.8980	0.8660	0.8230	0.8682	0.8860	0.8296	0.8780	0.8298	0.7883
	CPSPAN	0.7468	0.7455	0.6530	0.7770	0.7843	0.6924	0.8018	0.7410	0.6655	0.7482	0.7397	0.6542
	DVIMC	0.9830	0.9567	0.9627	0.9754	0.9386	0.9463	0.9588	**0.9060**	0.9115	0.9148	0.8363	0.8257
	RPCIC	0.9368	0.8735	0.8649	0.8742	0.7939	0.7552	0.7678	0.7049	0.6168	0.5714	0.5505	0.3977
	Mask-IMvC	0.9814	0.9548	0.9589	0.9700	0.9275	0.9346	0.9238	0.8551	0.8281	0.8837	0.8008	0.7361
	HSACC	0.9770	0.9524	0.9512	0.8870	0.9325	0.8694	0.7930	0.8645	0.7289	0.7778	0.8171	0.7340
	SCVT	0.9872	0.9650	0.9718	0.9092	0.8434	0.7969	0.9288	0.8611	0.8473	0.6148	0.6121	0.4398
	GHICMC	0.9864	0.9627	0.9701	0.9746	0.9346	0.9445	0.9558	0.8941	0.9046	0.8640	0.7876	0.7539
	HCCIMC	**0.9924**	**0.9783**	**0.9832**	**0.9796**	**0.9456**	**0.9553**	**0.9614**	0.9058	**0.9165**	**0.9454**	**0.8757**	**0.8832**
MSRC-V1	DCP	0.7952	0.7529	0.6637	0.4048	0.4483	0.3078	0.3143	0.2251	0.1239	0.2381	0.0796	0.0072
	CPSPAN	0.7524	0.7345	0.6468	0.7619	0.7131	0.6374	0.8810	0.7891	0.7571	0.7190	0.6844	0.5812
	DVIMC	0.8429	0.7350	0.6750	0.7286	0.6311	0.5486	0.6429	0.5607	0.4445	0.5714	0.4875	0.3339
	RPCIC	0.7143	0.6407	0.5657	0.5571	0.4642	0.3354	0.4000	0.2603	0.1461	0.2857	0.1453	0.0529
	Mask-IMvC	0.8200	0.8536	0.7599	0.6481	0.6421	0.4918	0.7067	0.6628	0.5272	0.6014	0.4848	0.3426
	HSACC	0.8190	0.7796	0.7134	0.3048	0.3304	0.1704	0.2190	0.0941	0.0016	0.2238	0.0607	0.0032
	SCVT	0.8381	0.8344	0.7507	0.8429	0.8153	0.7521	0.7905	0.7269	0.6529	0.7905	0.7357	0.6417
	GHICMC	**0.9543**	**0.9087**	**0.8937**	0.9372	**0.8766**	0.8514	0.9105	0.8246	0.8046	0.8448	0.7518	0.7101
	HCCIMC	0.9476	0.8995	0.8782	**0.9381**	0.8758	**0.8604**	**0.9143**	**0.8373**	**0.8109**	**0.9000**	**0.8179**	**0.7855**

shown in Figure 5(b), the clustering performance improves as K_n increases under low missing rates, while larger fluctuations are observed under high missing rates. The number of neighbors K_n used to construct the nearest-neighbor graph is selected from $\{3, 5, 7, 10\}$ for each dataset. The sensitivity analyses of K_a, λ, and γ are in the supplementary material.

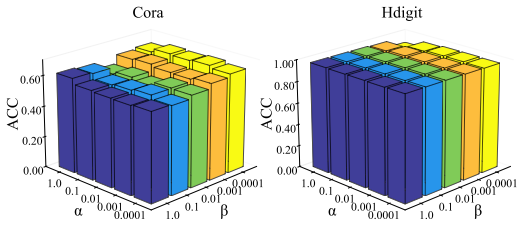


Figure 4: The parameter sensitivity analysis of α and β on two representative datasets with missing rate of 0.5.

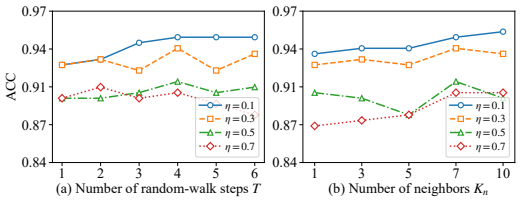


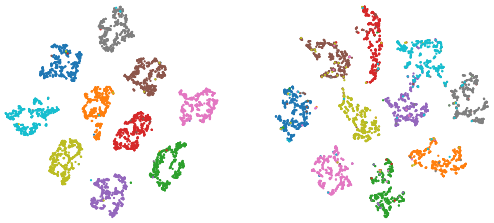
Figure 5: The sensitivity analysis of the random-walk steps T and the number of neighbors K_n on the MSRC-V1 dataset.

Table 3: Ablation studies on MSRC-V1 and Cora under different missing rates.

Components		MSRC-V1				Cora			
Completion	Contrastive	0.1	0.3	0.5	0.7	0.1	0.3	0.5	0.7
KNN	–	0.7857	0.7190	0.6857	0.6190	0.4978	0.4487	0.4287	0.4202
HCCM	–	0.7333	0.8048	0.7762	0.6429	0.6469	0.6086	0.5218	0.4716
KNN	MNCCL	0.9238	0.9048	0.9048	0.8857	0.6894	0.5764	0.6377	0.5421
HCCM	CL	0.9286	0.9333	0.7762	0.7000	0.6909	0.6768	0.6244	0.5705
HCCM	MNCCL	**0.9476**	**0.9381**	**0.9143**	**0.9000**	**0.7053**	**0.6869**	**0.6588**	**0.6285**

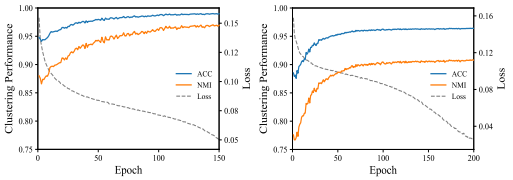
4.5 Ablation Study

To evaluate the effectiveness of the proposed HCCM and MNCCL modules, we conduct ablation studies by removing each module or replacing it with its corresponding baseline strategy. Specifically, HCCM is replaced with KNN-based completion, while MNCCL is replaced with conventional contrastive learning (CL). Table 3 reports the ablation results of HCCIMC variants under different missing rates. It can be observed that removing or replacing either HCCM or MNCCL leads to performance degradation, demonstrating the effectiveness of both modules. More experimental results are provided in the supplementary material.



(a) missing rate of 0.1 (b) missing rate of 0.7

Figure 6: The visualization results on the MNIST-USPS dataset.



(a) missing rate of 0.1 (b) missing rate of 0.7

Figure 7: The convergence analysis on the Hdigit dataset.

4.6 Visualization and Convergence Analysis

In order to visually investigate the effectiveness of the proposed HCCIMC, we use t-SNE to visualize the clustering results on the MNIST-USPS dataset, as shown in Figure 6. Our model can separate the feature representations of the different categories effectively. As the missing rate increases, the clustering results become less satisfactory than those under lower missing rates, but clear clustering structures can still be observed. This indicates that our method remains robust even under high missing rate conditions. To verify the convergence of the proposed method, we conduct experiments on Hdigit dataset to observe the changes in loss values and clustering metrics over training epochs. As shown in Figure 7, we can clearly observe that ACC and NMI steadily improve in the first few epochs, while the loss is quickly reduced. Afterwards, as the number of epochs increases, ACC and NMI continue to improve steadily and eventually converge to stable values.

5 Conclusion

In this paper, we propose HCCIMC, a high-order consensus completion method for incomplete contrastive multi-view clustering. To address unreliable missing-view completion and false positive pairs, HCCIMC introduces two complementary modules. Specifically, HCCM constructs view-specific bipartite graphs and aggregates their affinity structures into a global consensus graph, where multi-order random walks recover missing-view representations. MNCCL combines multi-order neighborhood relations with cross-view consistency to select reliable positive pairs, reducing the influence of noisy positives. Extensive experiments on four benchmark datasets demonstrate that HCCIMC achieves superior clustering performance over existing IMVC methods.

AI USE STATEMENT

In this work, generative AI tools were used to assist with language polishing, manuscript formatting, literature retrieval, and code optimization. We did not use generative AI tools to generate experimental results, fabricate data, or make final scientific claims, and other required-disclosure tasks that are not relevant to this work are not applicable. All AI-assisted content, including code-related modifications, was carefully reviewed, tested, and verified by the authors for correctness, consistency, and originality. The authors take full responsibility for the final content of this work, including any text, code, or artifacts produced with the assistance of generative AI.

REFERENCES

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*, volume 38, pp. 11221–11229, 2024.

Guoqing Chao, Kaixin Xu, Xijiong Xie, and Yongyong Chen. Global graph propagation with hierarchical information transfer for incomplete contrastive multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 15713–15721, 2025.

MS Chen, JQ Lin, XL Li, BY Liu, CD Wang, Dong Huang, and JH Lai. Representation learning in multi-view clustering: A literature review, data sci. *Eng*, 7(3):225–241, 2022.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pp. 1597–1607. PmLR, 2020.

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 Computer Vision and Pattern Recognition Conference*, pp. 5071–5081, 2025.

Xiaojian Ding, Lin Zhao, Xian Li, and Xiaoying Zhu. Incomplete multi-view clustering via hierarchical semantic alignment and cooperative completion. *Advances in Neural Information Processing Systems*, 38:52175–52195, 2026.

Yu Ding, Katsuya Hotta, Chunzhi Gu, Ao Li, Jun Yu, and Chao Zhang. Learning to discriminate while contrasting: Combating false negative pairs with coupled contrastive learning for incomplete multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 37(10): 6046–6060, 2025. doi: 10.1109/TKDE.2025.3592126.

Zhibin Dong, Dayu Hu, Jiaqi Jin, Siwei Wang, Xinwang Liu, and En Zhu. Selective cross-view topology for deep incomplete multi-view clustering. *IEEE Transactions on Image Processing*, 34:4792–4805, 2025a. doi: 10.1109/TIP.2025.3587586.

Zhibin Dong, Jiaqi Jin, Yuyang Xiao, Bin Xiao, Siwei Wang, Xinwang Liu, and En Zhu. Subgraph propagation and contrastive calibration for incomplete multiview data clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 36(2):3218–3230, 2025b. doi: 10.1109/TNNLS.2024.3350671.

Yiming Du, Ziyu Wang, Jian Li, Rui Ning, and Lusi Li. Deep incomplete multi-view clustering via hierarchical imputation and alignment. In *AAAI Conference on Artificial Intelligence*, 2026.

Si-Guo Fang, Dong Huang, Xiao-Sha Cai, Chang-Dong Wang, Chaobo He, and Yong Tang. Efficient multi-view clustering via unified and discrete bipartite graph learning. *IEEE Transactions on Neural Networks and Learning Systems*, 35(8):11436–11447, 2024. doi: 10.1109/TNNLS.2023.3261460.

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. doi: 10.1109/TKDE.2023.3270311.

Wei Feng, Guoshuai Sheng, Qianqian Wang, Quanxue Gao, Zhiqiang Tao, and Bo Dong. Partial multi-view clustering via self-supervised network. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, pp. 11988–11995, 2024.

Jun Guo and Jiahui Ye. Anchors bring ease: an embarrassingly simple approach to partial multi-view clustering. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence*, AAAI 19. AAAI Press, 2019. ISBN 978-1-57735-809-1. doi: 10.1609/aaai.v33i01.3301118.

Binqiang Huang, Zhijie Huang, Shoujie Lan, Qinghai Zheng, and Yuanlong Yu. Incomplete multi-view clustering via inference and evaluation. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8180–8184, 2024. doi: 10.1109/ICASSP48485.2024.10448378.

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*, pp. 11600–11609, 2023.

Jiaqi Jin, Siwei Wang, Zhibin Dong, Xihong Yang, Xinwang Liu, En Zhu, and Kunlun He. Deep incomplete multi-view clustering with distribution dual-consistency recovery guidance. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 1016–1026, October 2025.

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*, IJCAI '23, 2023. ISBN 978-1-956792-03-4. doi: 10.24963/ijcai.2023/435.

Yingming Li, Ming Yang, and Zhongfei Zhang. A survey of multi-view representation learning. *IEEE Transactions on Knowledge and Data Engineering*, 31(10):1863–1883, 2019. doi: 10.1109/TKDE.2018.2872063.

Zhenglai Li, Chang Tang, Xiao Zheng, Xinwang Liu, Wei Zhang, and En Zhu. High-order correlation preserved incomplete multi-view subspace clustering. *IEEE Transactions on Image Processing*, 31:2067–2080, 2022. doi: 10.1109/TIP.2022.3147046.

Zhenglai Li, Yuqi Shi, Xiao He, and Chang Tang. Mask-informed deep contrastive incomplete multi-view clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(12):11822–11833, 2025. doi: 10.1109/TCSVT.2025.3585600.

Yijie Lin, Yuanbiao Gou, Zitao Liu, Boyun Li, Jiancheng Lv, and Xi Peng. Completer: Incomplete multi-view clustering via contrastive prediction. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11169–11178, 2021. doi: 10.1109/CVPR46437.2021.01102.

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. doi: 10.1109/TPAMI.2022.3197238.

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, 2023.

Yiding Lu, Yijie Lin, Mouxing Yang, Dezhong Peng, Peng Hu, and Xi Peng. Decoupled contrastive multi-view clustering with high-order random walks. *Thirty-Eighth AAAI Conference on Artificial Intelligence*, 2024.

Erlin Pan and Zhao Kang. High-order multi-view clustering for generic data. *Information Fusion*, 100:101947, 2023. ISSN 1566-2535. doi: https://doi.org/10.1016/j.inffus.2023.101947.

Xi Peng, Zhenyu Huang, Jiancheng Lv, Hongyuan Zhu, and Joey Tianyi Zhou. COMIC: Multi-view clustering without parameter selection. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 5092–5101, Long Beach, California, USA, 09–15 Jun 2019. PMLR.

Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, Pierre-Antoine Manzagol, and Léon Bottou. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of machine learning research*, 11(12), 2010.

Qianqian Wang, Zhengming Ding, Zhiqiang Tao, Quanxue Gao, and Yun Fu. Generative partial multi-view clustering with adaptive fusion and cycle consistency. *IEEE Transactions on Image Processing*, 30:1771–1783, 2021a. doi: 10.1109/TIP.2020.3048626.

Qianqian Wang, Huanhuan Lian, Gan Sun, Quanxue Gao, and Licheng Jiao. icmsc: Incomplete cross-modal subspace clustering. *IEEE Transactions on Image Processing*, 30:305–317, 2021b. doi: 10.1109/TIP.2020.3036717.

Yiming Wang, Dongxia Chang, Zhiqiang Fu, Jie Wen, and Yao Zhao. Incomplete multiview clustering via cross-view relation transfer. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(1):367–378, 2023. doi: 10.1109/TCSVT.2022.3201822.

J. Winn and N. Jojic. Locus: learning object classes with unsupervised segmentation. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 1, pp. 756–763 Vol. 1, 2005. doi: 10.1109/ICCV.2005.148.

Danyang Wu, Zhenkun Yang, Jitao Lu, Jin Xu, Xiangmin Xu, and Feiping Nie. Ebmgc-gnf: Efficient balanced multi-view graph clustering via good neighbor fusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12):7878–7892, 2024a. doi: 10.1109/TPAMI.2024.3398220.

Song Wu, Yan Zheng, Yazhou Ren, Jing He, Xiaorong Pu, Shudong Huang, Zhifeng Hao, and Lifang He. Self-weighted contrastive fusion for deep multi-view clustering. *IEEE Transactions on Multimedia*, 26:9150–9162, 2024b. doi: 10.1109/TMM.2024.3387298.

Baili Xiao, Zhibin Dong, Ke Liang, Suyuan Liu, Siwei Wang, Tianrui Liu, Xingchen Hu, En Zhu, and Xinwang Liu. Easemvc: Efficient dual selection mechanism for deep multi-view clustering. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20716–20726, 2025. doi: 10.1109/CVPR52734.2025.01929.

Deng Xu, Chao Zhang, Cong Guo, Chunlin Chen, and Huaxiong Li. Fast incomplete multi-view clustering with adaptive similarity completion and reconstruction. In *Proceedings of the AAAI conference on artificial intelligence*, volume 39, pp. 21734–21742, 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)*, pp. 8761–8769, 2022a.

Jie Xu, Huayi Tang, Yazhou Ren, Liang Peng, Xiaofeng Zhu, and Lifang He. Multi-level feature learning for contrastive multi-view clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16051–16060, 2022b.

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.

Zhe Xue, Yawen Li, Zhongchao Guan, Wenling Li, Meiyu Liang, and Hai Zhou. Robust multi-graph contrastive network for incomplete multi-view clustering. *IEEE Transactions on Multimedia*, 27:2747–2759, 2025. doi: 10.1109/TMM.2023.3347639.

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*, MM '24, pp. 10402–10411, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400706868. doi: 10.1145/3664647.3681397.

Yuanyang Zhang, Yijie Lin, Weiqing Yan, Li Yao, Xinhang Wan, Guangyuan Li, Chao Zhang, Guanzhou Ke, and Jie Xu. Incomplete multi-view clustering via diffusion contrastive generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 22650–22658, 2025.

Deyin Zhuang, Jian Dai, Xingfeng Li, Xi Wu, Yuan Sun, and Zhenwen Ren. Robust graph contrastive learning for incomplete multi-view clustering. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, IJCAI 25, 2025. ISBN 978-1-956792-06-5.

A Appendix

A.1 Additional Experimental Settings

The network architecture, training strategy, and common hyperparameter search spaces are provided in the main paper. We report additional implementation details for reproducibility. For HCCM, view-specific anchors are constructed by applying K-means to the observed latent representations of each view. We use K-means++ initialization with $n_{\text{init}} = 1$, a maximum of 300 iterations, a tolerance of 10^{-4}, and a random seed of 42. The observed representations and anchors are ℓ_2-normalized, and the sample-to-anchor assignments are obtained using a row-wise Softmax over their cosine similarities with a temperature of 0.07. When constructing the view-specific affinity matrices, self-connections are removed and only the top-10 affinities in each row are retained. For the observation-constrained calibration, the ridge regularization coefficient, entropy regularization coefficient, and lower bound of the propagation weights are set to 0.01, 0.5, and 0.1, respectively. The propagation weights are optimized for 60 iterations using Adam with a learning rate of 0.2. The whole learning process of HCCIMC is summarized in Algorithm 1. All other implementation details are provided in the main paper.

A.2 Construction of Incomplete Multi-view Data

The incomplete rate is defined as $\eta = (N - N_c)/N$, where N and N_c denote the numbers of all samples and complete samples, respectively. Given η, we randomly select $\lfloor \eta N \rfloor$ samples as incomplete, while the remaining samples retain all views. For each incomplete sample i, we independently draw $r_{iv} \sim \mathcal{U}(0, 1)$ for each view and a shared threshold $q_i \sim \mathcal{U}(0, 1)$, and construct its view mask as $M_{iv} = \mathbb{I}[r_{iv} \geq q_i]$. The mask is resampled when all views are observed or all views are missing, thereby ensuring that each incomplete sample contains between 1 and $V - 1$ observed views. Under this sampling strategy, the number of missing views is randomly distributed over $\{1, \ldots, V - 1\}$. The mask seed is fixed to 10 for reproducibility.

A.3 Compared Methods

To comprehensively evaluate the proposed HCCIMC, we compare it with eight representative deep IMVC methods on four multi-view datasets. The compared methods are briefly introduced below.

- **DCP** (Lin et al., 2023): It maximizes the mutual information to learn consensus representation and minimizes the conditional entropy to recover missing data.
- **CPSPAN** (Jin et al., 2023): It aligns pair-observed samples and shifted prototypes across views, and imputes missing embeddings via cross-view structural transfer.
- **DVIMC** (Xu et al., 2024): It uses variational autoencoders and Product-of-Experts to learn a common representation from available views without imputing missing views.
- **RPCIC** (Yuan et al., 2024): It uses contrastive learning and robust prototype discrimination to alleviate noisy prototype correspondence, and completes missing data via prototype fusion.

Algorithm 1 HCCIMC

Input: Multi-view dataset $\{X^v\}_{v=1}^V$ with N samples and V views, number of clusters K, number of anchors K_a, maximum random-walk step number T, temperature parameter τ_c, parameters α and β, maximum iterations $epochs$.
Output: Clustering result $\mathbf{Y}$.
Initialize autoencoder parameters by minimizing $\mathcal{L}_{rec}$ in Eq. (3)
for $epoch = 1$ to $epochs$ **do**
    Calculate the loss $\mathcal{L}_{rec}$ by Eq. (3)
    Compute transition matrix $\mathbf{P}$ by Eq. (7)
    Compute propagation dictionary $\mathcal{T}^v$ by Eq. (9)
    Compute completion representation $\hat{z}^v$ by Eq. (11)
    Compute positive sample set $\mathcal{P}$ by Eq. (17)
    Calculate the loss $\mathcal{L}_{cl}$ by Eq. (19)
    Calculate the overall loss $\mathcal{L}$ by Eq. (20)
    Update through gradient descent to minimize $\mathcal{L}$
end for
Compute consensus representation $Z = \frac{1}{V}\sum_{v=1}^{V} \widehat{Z}^v$
Obtain the final clustering result $\mathbf{Y}$ by k-means
return $\mathbf{Y}$

- **Mask-IMVC** (Li et al., 2025): It learns a view-common representation via mask-informed fusion and weighted contrastive learning without imputing missing views.
- **SCVT** (Dong et al., 2025a): It constructs a view topology graph for missing-view recovery and employs graph-guided contrastive learning to obtain clustering-friendly representations.
- **GHICMC** (Chao et al., 2025): It uses view-specific GCNs and global graph propagation to impute missing representations.
- **HSACC** (Ding et al., 2026): It performs hierarchical semantic alignment with dynamic weighted fusion and uses cross-view inference to recover missing-view representations.

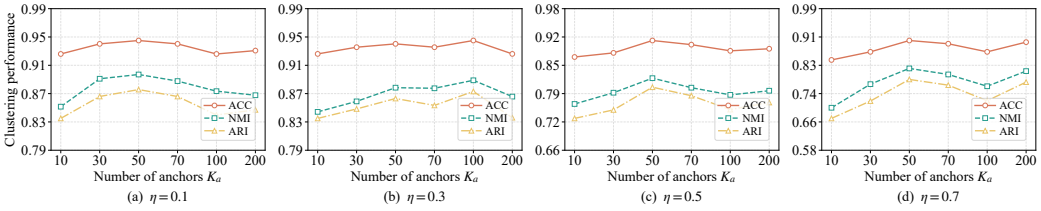


Figure 8: The effect of the number of anchors K_a on the MSRC-V1 dataset under different missing rates η.

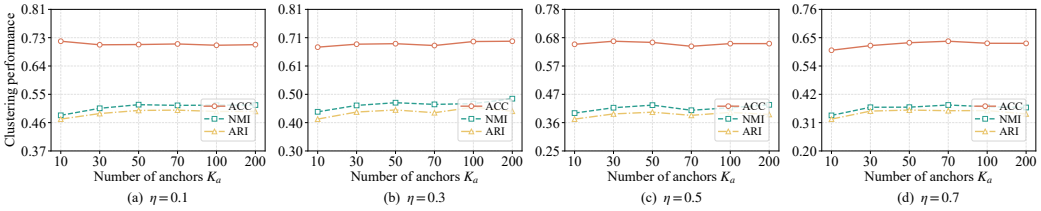


Figure 9: The effect of the number of anchors K_a on the Cora dataset under different missing rates η.

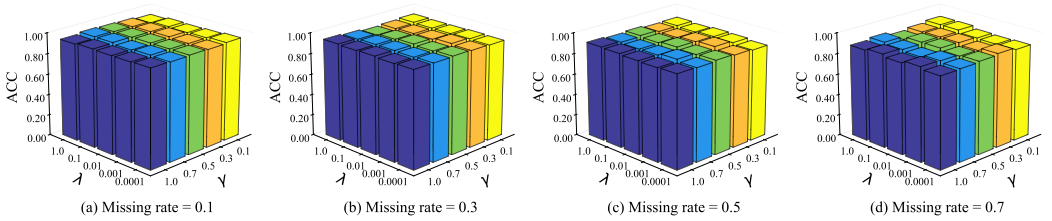


Figure 10: The effect of the regularization coefficients λ and γ on the MSRC-V1 dataset under different missing rates η.

A.4 Parameter Sensitivity Analysis

We conduct sensitivity analysis on the number of anchors K_a, the ridge regularization coefficient λ, and the entropy regularization coefficient γ. When performing sensitivity experiments for a single parameter, all other parameters are fixed at their optimal values. As shown in Figures 8 and 9, clustering performance on MSRC-V1 generally improves as K_a increases to 50 and then fluctuates slightly, whereas that on Cora remains relatively stable. These results demonstrate that the proposed method is robust to K_a over a broad range. We set $K_a = 50$ for all datasets. Figure 10 illustrates the sensitivity of HCCIMC to the ridge regularization coefficient λ and the entropy regularization coefficient γ. Within the selected ranges, the clustering performance remains relatively stable across most parameter combinations, although minor fluctuations are observed as the missing rate increases. This indicates that HCCIMC is not highly sensitive to either coefficient. Therefore, we adopt the moderate settings $\lambda = 0.01$ and $\gamma = 0.5$ in all experiments.

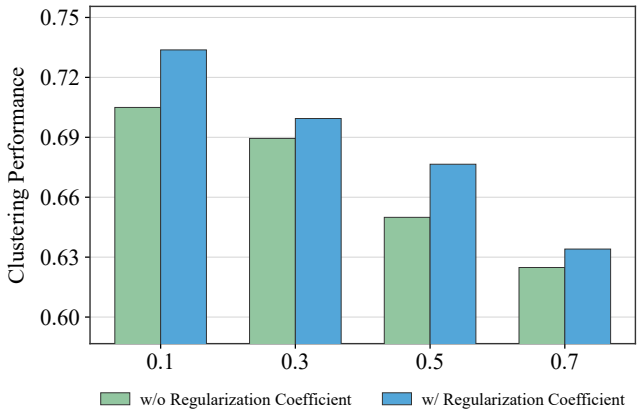


Figure 11: Ablation study of the regularization term on the Cora dataset under different missing rates.

A.5 Effectiveness of Regularization Terms

As shown in Figure 11, introducing the regularization terms λ and γ consistently improves clustering performance under different missing rates. Specifically, the two terms jointly regulate the magnitude and distribution of the propagation weights, thereby preventing unstable or overly concentrated solutions and enabling more reliable aggregation of multi-order information. These results demonstrate that the regularization terms effectively constrain the propagation weights and enhance the robustness of HCCIMC.

A.6 Computational Complexity Analysis

Let N, V, and d denote the number of samples, the number of views, and the latent feature dimension, respectively. Let D denote the maximum number of neurons in the hidden layers of the

autoencoders, K_a denote the number of anchors, T denote the maximum propagation order, and N_b denote the mini-batch size. The computational complexity of the view-specific representation learning module is $\mathcal{O}(NVD^2)$. In HCCM, anchor graph construction, consensus graph construction, and multi-order propagation require $\mathcal{O}(NVK_ad)$, $\mathcal{O}(VN^2K_a)$, and $\mathcal{O}(VTN^2d)$, respectively. The computational complexity of MNCCL is $\mathcal{O}(V^2NN_b(d+N_b))$. Since V, D, K_a, T, d, and N_b are relatively small, the overall computational complexity is approximately $\mathcal{O}(VN^2(K_a+Td))$, which is asymptotically equivalent to $\mathcal{O}(N^2)$.