STRUCTURE-AWARE CONDITIONAL DRIFT GENERATION FOR INCOMPLETE MULTI-VIEW CLUSTERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Incomplete multi-view clustering (IMVC) aims to discover latent structures from multi-view data with partially unavailable observations. However, existing methods often recover missing information through feature estimation, cross-view averaging, or neighborhood transfer, which may overlook the target-view distribution and weaken sample correspondence. Although recent approaches improve missing-view recovery by exploiting cross-view consistency or local relationships, they generally provide limited joint modeling of local geometry and cluster-level semantics, while the reliability of recovered representations is rarely considered during subsequent fusion. To tackle these issues, we propose Structure-Aware Conditional Drift Generation (SACDG) for IMVC. Specifically, SACDG takes shared semantics from observed views as a reliable cross-view reference and learns a structure-guided drift toward the target-view distribution, thereby preserving sample correspondence while adapting to target-specific geometry and cluster semantics. Additionally, pseudo-missing self-supervision estimates recovery reliability to reduce the influence of inaccurate completions during multi-view fusion. Finally, structural alignment and prototype-driven clustering provide consistent feedback for representation recovery and clustering. Extensive experiments on seven benchmark datasets under multiple missing rates demonstrate the effectiveness and robustness of SACDG compared with state-of-the-art IMVC methods.

1 INTRODUCTION

Multi-view clustering (MVC) seeks to exploit complementary and consistent information from multiple feature spaces or modalities to learn discriminative representations and clustering assignments. Deep MVC has substantially improved nonlinear representation learning by modeling semantic consistency, local topology, and cross-view interactions (Chen et al., 2023; Tan et al., 2023; Yan et al., 2023). Tensor correlation modeling and information bottlenecks further enhance scalability and compactness (Long et al., 2024; Yan et al., 2024), while learnable anchors, mixture-of-experts representations, adaptive fusion, and correlation-based permutations provide more flexible cross-view integration (Wang et al., 2025; Zhang et al., 2025c; Zhao et al., 2026; Eisenberg et al., 2025). These advances highlight the importance of jointly exploiting cross-view semantics and structures. In particular, preserving semantic agreement while retaining view-specific structural information is crucial for stable clustering across heterogeneous views under complex missing scenarios.

However, real-world data are often incomplete, giving rise to incomplete multi-view clustering (IMVC). Existing methods recover missing information through consensus semantics, anchor alignment, similarity fusion, or kernel coupling (Dai et al., 2025; Qin et al., 2025; Jiang et al., 2025a; Wu et al., 2025), while high-order and deep approaches exploit graph completion, low-rank recovery, reliable imputation, or similarity reconstruction (Zhao et al., 2023; Zhang et al., 2023; Tang & Liu, 2022; Xu et al., 2025). Despite these advances, most methods directly estimate unavailable representations, although observed views provide semantics differing from the target-view distribution. Copying or averaging observed information may preserve correspondence but suppress target-view specificity, whereas unconstrained generation may weaken the semantic link to the original sample.

This observation motivates a different perspective in which missing-view recovery should learn how shared semantics gradually drift toward the target-view distribution rather than reconstructing the target independently. Such drift must preserve sample correspondence while effectively adapting

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.

to target-view local geometry and cluster semantics. This is particularly important because missing views simultaneously disrupt feature values, neighborhood relations, and cluster-level structure. Simple averaging or neighbor-based recovery may over-smooth representations, while imputation errors can be amplified during fusion and self-training, especially at high missing rates. Recent studies alleviate distribution bias and mismatched noise (Wang et al., 2026), or strengthen cross-view correspondence through cyclic permutation in a VAE latent space (Gao & Pu, 2025). Nevertheless, explicitly learning a structure-guided target-view drift from a reliable cross-view reference, while estimating the reliability of each recovered representation, remains insufficiently explored.

To address these issues, we propose Structure-Aware Conditional Drift Generation (SACDG) for IMVC. SACDG takes a cross-view semantic anchor as the starting point for missing-view generation and learns a target-view-specific conditional drift under local graph context and shared prototype semantics. In this way, recovery retains reliable sample-level semantics while progressively adapting them toward the target-view distribution in a controlled sample-specific manner. Generation confidence is further estimated to calibrate recovered representations and guide reliability-aware fusion, while graph propagation and prototype-driven clustering provide structural feedback for joint optimization. Thus, missing-view recovery and clustering mutually refine each other within a unified structure-aware learning process. The main contributions are summarized as follows:

- To the best of our knowledge, we make the first attempt to integrate conditional drift learning into missing-view imputation for IMVC. Instead of directly estimating missing representations, we formulate recovery as learning a target-view-specific drift from a reliable cross-view semantic anchor, with graph context and prototype semantics jointly guiding the shift toward the target-view distribution in a controlled progressive manner.
- We introduce generation-reliability estimation and reliability-aware fusion, using pseudo-missing self-supervision to calibrate confidence and further suppress the propagation of unreliable completion errors during subsequent multi-view fusion and clustering.
- We develop a joint framework that integrates graph-structure alignment with prototype-driven clustering, enabling iterative interaction among representation learning, missing-view recovery, and clustering optimization through progressive structural feedback.

2 Related Work

2.1 Deep Incomplete Multi-View Clustering

Deep incomplete multi-view clustering jointly learns view-specific representations, recovers missing information, and optimizes clustering structures within a unified and mutually reinforcing learning process. Early deep methods improve recovery by selecting reliable semantic neighbors and exploiting high-confidence cross-view consistency (Tang & Liu, 2022; Chao et al., 2024), while variational representation learning and latent-graph-guided imputation further strengthen shared semantic modeling and preserve local relationships across incomplete views (Xu et al., 2024; Pu et al., 2024). Recent studies have moved toward more challenging incomplete settings, focusing on robust representation learning under irregular missing patterns and structurally consistent recovery through distribution-aware modeling and graph-based information propagation (Jiang et al., 2025b; Jin et al., 2025; Du et al., 2026; Chen et al., 2026; Zhang et al., 2026; Wang et al., 2026). Despite these advances, most existing methods still rely on feature estimation or cross-view alignment for missing-view recovery, with limited attention to structural guidance during generation. Meanwhile, the reliability of recovered representations is rarely considered in subsequent fusion or clustering, which may amplify completion errors under diverse missing patterns. These limitations motivate us to develop a structure-aware, reliability-aware generation framework for more robust IMVC.

2.2 Generative Imputation for Incomplete Multi-View Clustering

Generative imputation explicitly reconstructs unavailable information by learning conditional distributions or latent mappings from observed views. Variational and distribution-consistent approaches recover missing views by modeling latent completeness or modality distributions (Cai et al., 2023; Wang et al., 2023), while diffusion-based methods improve generation quality and semantic consistency under incomplete observations (Wen et al., 2024; Zhang et al., 2025b). Cross-view generation

has also been realized through latent-variable permutation and cyclic posterior constraints (Gao & Pu, 2025). Although these methods demonstrate the potential of generative recovery, they rarely use local graph structure and current clustering prototypes together as generation conditions or explicitly calibrate the reliability of individual generated representations. SACDG addresses these limitations through structure-aware conditional drift generation and confidence-guided fusion.

3 METHOD

3.1 PROBLEM DEFINITION

Given incomplete multi-view data $\{X^v\}_{v=1}^V$ with N samples and V views, the v-th view is $X^v \in \mathbb{R}^{N\times d_v}$, where d_v is its feature dimension. A binary mask $M \in \{0,1\}^{N\times V}$ records observation status, with $M_i^v = 1$ for an observed entry and $M_i^v = 0$ otherwise. For sample i, the observed- and missing-view sets are $\Omega_i = \{v \mid M_i^v = 1\}$ and $\overline{\Omega}_i = \{v \mid M_i^v = 0\}$, with $|\Omega_i| \geq 1$. Given K clusters, IMVC aims to learn a unified representation and predict labels Y without ground-truth supervision for robust clustering under incomplete observations (Tang & Liu, 2022; Xu et al., 2022).

3.2 VIEW REPRESENTATION AND STRUCTURE PRESERVATION

SACDG first maps heterogeneous views into a shared d-dimensional latent space to effectively reduce instability from heterogeneous feature scales and dimensions while retaining view-specific information. The latent representation H^v and reconstruction $\hat{X}^v$ of view v are obtained by:

$$H^v = E^v(X^v), \qquad \hat{X}^v = D^v(H^v), \tag{1}$$

where $E^v(\cdot)$ and $D^v(\cdot)$ are the view-specific encoder and decoder, $H^v = [h_1^v, \ldots, h_N^v]^\top \in \mathbb{R}^{N\times d}$ is the latent matrix, h_i^v denotes the latent code of sample i, and $\hat{X}^v$ is the reconstructed view. This branch directly preserves observed view-specific content before missing-view generation.

Reconstruction uses observed samples. Let $O_v = \{i \mid M_i^v = 1\}$ index observations in view v. To prevent noisy residuals dominating learning, the robust feature reconstruction loss L_{feat}^v is:

$$L_{\text{feat}}^v = \frac{1}{|O_v| d_v} \sum_{i\in O_v} \sum_{j=1}^{d_v} \phi_\delta(x_{ij}^v - \hat{x}_{ij}^v), \tag{2}$$

where x_{ij}^v and $\hat{x}_{ij}^v$ are the original and reconstructed feature values. The robust penalty $\phi_\delta(e)$ is:

$$\phi_\delta(e) = \begin{cases} \dfrac{e^2}{2\delta}, & |e| < \delta, \\ |e| - \dfrac{\delta}{2}, & |e| \geq \delta, \end{cases} \tag{3}$$

where e is the reconstruction residual and $\delta > 0$ controls the transition between quadratic and linear regions. Based on empirical evaluation, we set $\delta = 1.0$ throughout all experiments, yielding smooth optimization for small errors while more effectively limiting the influence of large residuals.

Feature reconstruction alone ignores local relations. SACDG therefore constrains the neighborhood structure of each observed view. Let $\mathcal{E}_v$ be the Euclidean KNN edge set in view v. Empirical evaluation selects $k_g = 12$, which is fixed for KNN graphs during training and inference. For an edge $(i, j) \in \mathcal{E}_v$, let s_{ij}^v and $\hat{s}_{ij}^v$ denote cosine similarities between the original and reconstructed feature vectors:

$$s_{ij}^v = \frac{(x_i^v)^\top x_j^v}{\|x_i^v\|_2 \|x_j^v\|_2}, \qquad \hat{s}_{ij}^v = \frac{(\hat{x}_i^v)^\top \hat{x}_j^v}{\|\hat{x}_i^v\|_2 \|\hat{x}_j^v\|_2}, \tag{4}$$

where x_i^v and $\hat{x}_i^v$ are the original and reconstructed feature vectors, and $\|\cdot\|_2$ denotes the Euclidean (ℓ_2) norm. Matching these similarities on local edges gives the structural reconstruction loss L_{str}^v:

$$L_{\text{str}}^v = \frac{1}{|\mathcal{E}_v|} \sum_{(i,j)\in\mathcal{E}_v} \phi_\delta(s_{ij}^v - \hat{s}_{ij}^v). \tag{5}$$

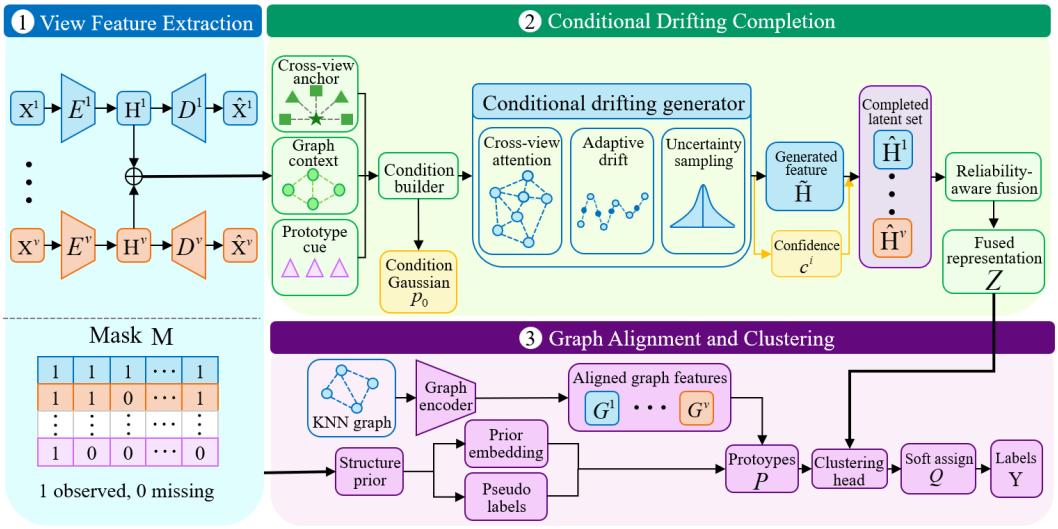


Figure 1: Overall framework of SACDG. (1) View feature extraction learns representations from multi-view data. (2) Conditional drifting completion recovers missing representations with structure-aware conditions and performs reliability-aware fusion. (3) Graph alignment and clustering refine representations and assignments through graph propagation and prototype-driven learning.

Together, L^v_{feat} preserves observed content and L^v_{str} preserves local geometry during pretraining. Their combination naturally defines the reconstruction objective L_{rec} used for warm-up:

$$L_{\text{rec}} = \frac{1}{V}\sum_{v=1}^{V}(L^v_{\text{feat}} + L^v_{\text{str}}). \tag{6}$$

3.3 Structure-Aware Conditional Drift Completion

Copying or averaging observed views may ignore the target-view distribution and over-smooth sample-specific information. SACDG instead models recovery as a controlled drift from shared semantics toward the target view. This formulation preserves a reliable semantic starting point while allowing more effective target-specific adaptation under structural and clustering guidance.

Definition 3.1 *Conditional drift learns a target-view residual from a cross-view anchor under structural, prototype-semantic, and stochastic conditions, moving shared semantics toward the target-view distribution through controlled adaptation while retaining the anchor as a stable reference.*

Definition 3.2 *For sample i and target view v, the generative condition comprises a cross-view anchor a^v_i, graph context s^v_i, and a prototype cue from P, so that generation reliably retains sample correspondence, local geometry, and clustering semantics during target-view recovery.*

Following Definition 3.2, SACDG first builds the cross-view anchor a^v_i from the observed representations of the same sample as:

$$a^v_i = \frac{1}{|\Omega_i|}\sum_{u\in\Omega_i} h^u_i. \tag{7}$$

The anchor $a^v_i \in \mathbb{R}^d$ preserves sample correspondence by aggregating the observed representations.

However, sample-level semantics alone cannot adequately describe the local manifold around the missing entry. SACDG therefore constructs the graph context s^v_i by aggregating neighbors from the observed source views:

$$s^v_i = \frac{1}{|\Omega_i|}\sum_{u\in\Omega_i}\frac{1}{|N^u_i|}\sum_{j\in N^u_i} h^u_j, \tag{8}$$

where N_i^u is the Euclidean KNN neighborhood of sample i in source view u and $|N_i^u| = k_g = 12$. The setting $k_g = 12$ is empirically selected through validation experiments. The context $s_i^v \in \mathbb{R}^d$ summarizes its local region and promotes neighborhood continuity in the generated representation.

To introduce cluster-level semantics, SACDG maintains prototypes $P = [P_1, \ldots, P_K] \in \mathbb{R}^{K \times d}$ and soft assignments $Q = [q_{ik}] \in [0,1]^{N \times K}$, where P_k represents cluster k and q_{ik} is the assignment probability. The cue $\sum_{k=1}^{K} q_{ik} P_k$ encourages consistency with the clustering structure.

The anchor, graph context, and prototype cue respectively constrain instance correspondence, local geometry, and cluster semantics. Given these conditions and random perturbation $\dot{q} \sim \mathcal{N}(0, I_d)$, the generated target-view representation $\tilde{h}_i^v$ is obtained through an anchor-relative conditional drift as:

$$\tilde{h}_i^v = a_i^v + f_\theta(a_i^v, s_i^v, \sum_{k=1}^{K} q_{ik} P_k, \dot{q}), \tag{9}$$

where $f_\theta(\cdot)$ is the drift generator, $\dot{q} \in \mathbb{R}^d$ is Gaussian perturbation, and I_d is the identity matrix. The residual form keeps a_i^v as a stable baseline while $f_\theta(\cdot)$ estimates the target-view displacement.

Because drift quality can vary across missing samples, SACDG should not trust every generated representation equally. The generator therefore outputs confidence $c^i \in [0,1]$ to quantify sample-specific recovery reliability. The completed representation $\hat{h}_i^v$ keeps the real latent code for observed entries and interpolates between the generated representation and anchor for missing entries:

$$\hat{h}_i^v = M_i^v h_i^v + (1 - M_i^v)[c^i \tilde{h}_i^v + (1 - c^i) a_i^v], \tag{10}$$

where larger c^i admits more information from $\tilde{h}_i^v$, while smaller c^i shifts the completion toward a_i^v to further suppress unreliable generation for more reliable downstream multi-view fusion.

Since true missing views are unavailable, pseudo-missing self-supervision is used. Let S contain temporarily masked observed sample–view pairs used as latent targets. The prediction loss L_{pred} is:

$$L_{\text{pred}} = \frac{1}{|S|} \sum_{(i,v) \in S} [\frac{1}{d} \sum_{j=1}^{d} \phi_\delta(\tilde{h}_{ij}^v - h_{ij}^v) + \lambda_{\cos}(1 - \cos(\tilde{h}_i^v, h_i^v))], \tag{11}$$

where $\lambda_{\cos} \geq 0$ balances component-wise recovery and cosine-direction consistency. The two terms jointly preserve target values and latent direction for more stable recovery under missingness.

To make confidence reflect actual recovery quality rather than an unconstrained score, the same pseudo-missing entries supervise confidence estimation. SACDG converts the recovery error into the soft reliability target ρ_i^v for subsequent reliability-aware completion and fusion:

$$\rho_i^v = \exp\left(-\|\tilde{h}_i^v - h_i^v\|_2^2 / \kappa_v\right), \tag{12}$$

where $\rho_i^v \in (0,1]$ is the recovery-reliability target and $\kappa_v > 0$ is a view-specific error scale. Smaller recovery errors yield larger reliability for subsequent fusion. The confidence loss L_{conf} is:

$$L_{\text{conf}} = -\frac{1}{|S|} \sum_{(i,v) \in S} [\rho_i^v \log c^i + (1 - \rho_i^v) \log(1 - c^i)]. \tag{13}$$

The overall completion objective L_{imp} is:

$$L_{\text{imp}} = L_{\text{pred}} + L_{\text{conf}}, \tag{14}$$

where L_{pred} supervises recovery and L_{conf} calibrates confidence, limiting unreliable completion.

3.4 Graph Structure Alignment and Reliability-Aware Fusion

After completion, point-wise recovery may still inadvertently alter local neighborhood relations between observed and generated entries. Let $\hat{H}^v = [\hat{h}_1^v, \ldots, \hat{h}_N^v]^\top$ denote the completed representation matrix of view v. SACDG therefore builds a KNN graph on $\hat{H}^v$ using the same $k_g = 12$

and employs a shared residual graph encoder to obtain the structure-enhanced representation g_i^v of sample i in view v, while preserving local sample information through neighborhood propagation:

$$g_i^v = \eta \hat{h}_i^v + (1-\eta)\sigma(W_s \hat{h}_i^v + W_n \frac{1}{|N_i^v|} \sum_{j \in N_i^v} \hat{h}_j^v), \tag{15}$$

where $\eta \in [0,1]$ controls residual retention, W_s and W_n are shared self and neighbor transformations, $\sigma(\cdot)$ is GELU, and N_i^v is the KNN neighborhood in completed view v. Residual retention and neighbor aggregation preserve sample information while propagating local structure.

Since completed views have unequal reliability across different missing samples, direct averaging may reintroduce completion noise during fusion. SACDG therefore defines the normalized reliability-aware view weight ω_i^v from the observation mask and generation confidence as:

$$\omega_i^v = \frac{M_i^v + (1-M_i^v)c^i}{\sum_{u=1}^{V}(M_i^u + (1-M_i^u)c^i)}, \qquad \sum_{v=1}^{V} \omega_i^v = 1, \tag{16}$$

where ω_i^v is the normalized contribution of view v to sample i. Observed views receive unit unnormalized weight, while missing views are weighted by c^i. The unified representation z_i is then:

$$z_i = \sum_{v=1}^{V} \omega_i^v g_i^v. \tag{17}$$

Accordingly, $Z = [z_1, \ldots, z_N]^\top$ denotes the resulting unified representation matrix. Reliability-aware weighting reduces the influence of inaccurate recovery on subsequent clustering.

3.5 Prototype-Driven Clustering and Joint Alignment

The unified matrix Z updates cluster assignments and prototypes during subsequent joint optimization stages. By comparing z_i with the shared prototypes, the soft assignment q_{ik} of sample i to cluster k is computed as:

$$q_{ik} = \frac{\exp(z_i^\top P_k/\tau)}{\sum_{l=1}^{K}\exp(z_i^\top P_l/\tau)}, \tag{18}$$

where $\tau > 0$ is the softmax temperature. The normalized assignments are used for clustering and provide semantic feedback to the drift generator. Each prototype is computed as an assignment-weighted soft cluster center:

$$P_k = \frac{\sum_{i=1}^{N} q_{ik} z_i}{\sum_{i=1}^{N} q_{ik}}, \tag{19}$$

where P_k is the soft center of cluster k, computed from assignment-weighted fused representations. The resulting prototypes provide cluster semantics to the drift generator for iterative refinement.

During joint optimization, consistency alone is insufficient because fusion must also preserve neighborhoods and sharpen cluster structure. We therefore coordinate three key complementary objectives. First, the reliability-weighted fusion-consistency loss L_{fus} aligns each graph representation with the fused space:

$$L_{\text{fus}} = \frac{1}{N}\sum_{i=1}^{N}\sum_{v=1}^{V} \omega_i^v [1 - \cos(g_i^v, z_i)]. \tag{20}$$

The reliability weights give observed and confident views stronger alignment signals throughout the subsequent fusion and clustering process while down-weighting unreliable completions.

Second, the topology-preservation loss L_{top} compares edge similarities in the view-specific graph space and the fused space, explicitly constraining the fused representation to more faithfully retain neighborhood relationships encoded by individual completed views during fusion:

$$L_{\text{top}} = \frac{1}{V}\sum_{v=1}^{V}\frac{1}{|\mathcal{E}_v|}\sum_{(i,j)\in\mathcal{E}_v} [\cos(g_i^v, g_j^v) - \cos(z_i, z_j)]^2. \tag{21}$$

This discrepancy discourages fusion from distorting stable local neighborhoods during integration.

Finally, high-confidence assignments are used for prototype-driven self-training. Let $C = \{i \mid \max_k q_{ik} \geq \xi\}$ be the selected high-confidence sample set, where $\xi \in (0, 1)$ is the selection threshold, and let $\hat{y}_i = \arg\max_k q_{ik}$ be the pseudo-label of sample i. The clustering loss L_{clu} is:

$$L_{\text{clu}} = -\frac{1}{|C|} \sum_{i \in C} \log q_{i\hat{y}_i}. \tag{22}$$

Using only high-confidence samples reduces pseudo-label noise while encouraging compact prototype-centered clusters. The combined objective L_{FCA} integrates fusion, topology preservation, and clustering as:

$$L_{\text{FCA}} = L_{\text{fus}} + L_{\text{top}} + L_{\text{clu}}. \tag{23}$$

3.6 Overall Objective and Training Strategy

SACDG combines view reconstruction, conditional drift completion, and graph-prototype clustering through the objective L:

$$L = L_{\text{rec}} + \lambda_{imp} L_{\text{imp}} + \lambda_{FCA} L_{\text{FCA}}, \tag{24}$$

where λ_{imp} and λ_{FCA} balance completion and joint fusion, clustering, and alignment relative to reconstruction. This objective links representation learning, drift recovery, and clustering feedback in a unified optimization process. Training first warms up the encoders and decoders with L_{rec} and then jointly optimizes modules with L. Algorithmic flow and complexity are provided in Appendix B.4.

4 Experiments

4.1 Experimental Setup

1) Datasets: We evaluate SACDG on seven benchmarks: **Synthetic3D** (Dong et al., 2025), **HW** (Zhang et al., 2025a), **MNIST4**, **MNIST-Fea** (Zhang et al., 2025b), **CUB** (Gao & Pu, 2025), **Scene15** (Li et al., 2025), and **Reuters** (Jin et al., 2025). Missing views are generated by random masking at the prescribed rates while ensuring that each sample retains at least one observed view. Detailed dataset descriptions are provided in Appendix C.1, covering diverse application scenarios.

2) Compared Methods: We compare SACDG with eight recent IMVC baselines: **BRIDGE** (Jiang et al., 2025b), **SCVT** (Dong et al., 2025), **FreeCSL** (Dai et al., 2025), **GAVIM** (Chen et al., 2026), **IMVC-NCD** (Guo et al., 2026), **DGIMVCM** (Zhang et al., 2026), **MVP** (Gao & Pu, 2025), and **CAMEL** (Gu et al., 2026). Their main modeling characteristics are summarized in Appendix C.1.

3) Evaluation Metrics: We evaluate clustering performance using Clustering Accuracy (ACC), Normalized Mutual Information (NMI), and Adjusted Rand Index (ARI), which measure assignment correctness, shared information between predicted and ground-truth partitions, and agreement between sample pairs after chance correction, respectively. Higher values indicate better performance.

4.2 Experimental Results

1) Experimental Results Analysis: Table 1 compares SACDG with eight representative baselines on seven datasets under four missing rates. SACDG achieves the best result in all settings, demonstrating consistent superiority across synthetic, visual, handwritten-digit, and document data. For instance, on Reuters, it obtains 61.83%, 38.51%, and 35.27% in ACC, NMI, and ARI at a missing rate of 0.1, while maintaining 55.47%, 32.80%, and 30.61% at 0.7. At the same severe missing rate, SACDG reaches 86.39% ACC on HW, outperforming the second-best DGIMVCM by 4.10 percentage points. These gains remain evident as missingness increases, validating the effectiveness of structure-aware conditional drift completion, graph-based structural modeling, and reliability-aware fusion in suppressing error propagation and preserving discriminative clustering structures.

2) Ablation Study: To examine the effect of each objective, we conduct ablation experiments on MNIST Fea, Scene15, and Reuters at a 10% missing rate. As shown in Table 2, L_{rec} in Eq.(6) is retained in all variants, while L_{imp} in Eq.(14), L_{FCA} in Eq.(23), or both are removed. The full model obtains the best ACC, NMI, and ARI on all three datasets. Compared with the strongest reduced variant, its ACC is higher by 7.88, 20.07, and 23.11 percentage points on MNIST Fea, Scene15, and

Table 1: Results (mean ± std, %) on seven datasets under four missing rates. Best and second-best results are highlighted in red and blue, respectively.

Dataset	Method	0.1			0.3			0.5			0.7		
		ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
Synthetic3D	BRIDGE (ICCV'25)	95.50 ± 1.00	82.32 ± 3.04	87.04 ± 2.51	93.45 ± 1.17	76.36 ± 3.27	81.16 ± 3.03	**90.58 ± 1.29**	**71.25 ± 3.22**	74.46 ± 3.87	**87.94 ± 1.19**	**63.67 ± 3.06**	**67.28 ± 2.96**
	SCVT (TIP'25)	**97.25 ± 0.11**	**88.19 ± 0.23**	**91.97 ± 0.19**	**94.50 ± 0.10**	75.21 ± 0.22	**82.84 ± 0.18**	89.55 ± 0.15	67.64 ± 0.33	**75.43 ± 0.21**	83.75 ± 0.16	60.03 ± 0.38	65.49 ± 0.15
	FreeCSL (CVPR'25)	96.17 ± 0.23	84.53 ± 0.85	88.92 ± 1.14	86.34 ± 0.26	74.04 ± 0.99	79.64 ± 1.02	80.55 ± 0.39	63.82 ± 0.86	60.70 ± 1.07	74.44 ± 0.35	46.35 ± 1.25	47.48 ± 1.32
	GAVIM (TKDE'26)	89.00 ± 0.82	80.15 ± 0.87	81.94 ± 0.89	84.58 ± 0.92	69.51 ± 1.30	70.61 ± 1.22	79.63 ± 0.69	59.32 ± 1.00	60.15 ± 1.12	76.57 ± 0.83	47.93 ± 1.26	50.85 ± 1.15
	IMVC-NCD (ICML'26)	94.58 ± 0.54	81.09 ± 0.91	84.77 ± 0.89	89.57 ± 0.44	70.94 ± 0.78	74.86 ± 1.30	85.64 ± 0.61	63.04 ± 1.07	65.91 ± 0.91	79.89 ± 0.48	52.80 ± 0.90	53.87 ± 1.09
	DGIMVCM (AAAI'26)	96.84 ± 0.13	86.90 ± 0.14	90.79 ± 0.17	92.53 ± 0.12	**77.20 ± 0.15**	81.57 ± 0.15	87.95 ± 0.13	68.62 ± 0.18	73.20 ± 0.18	82.95 ± 0.10	59.21 ± 0.21	62.13 ± 0.17
	MVP (ICLR'25)	95.25 ± 0.44	82.19 ± 0.35	86.46 ± 0.58	89.20 ± 0.59	72.82 ± 0.52	76.20 ± 0.59	85.38 ± 0.44	60.77 ± 0.40	63.80 ± 0.49	81.65 ± 0.40	54.20 ± 0.44	54.21 ± 0.48
	CAMEL (ICLR'26)	85.50 ± 0.54	71.13 ± 0.50	73.07 ± 0.53	82.13 ± 0.56	66.39 ± 0.57	67.42 ± 0.44	77.73 ± 0.67	61.16 ± 0.60	59.88 ± 0.54	74.02 ± 0.76	55.26 ± 0.50	54.06 ± 0.48
	SACDG (Ours)	**97.75 ± 0.78**	**89.45 ± 2.03**	**92.93 ± 1.99**	**95.00 ± 0.88**	**81.51 ± 1.49**	**83.93 ± 2.00**	**91.92 ± 0.79**	**74.21 ± 2.20**	**75.88 ± 2.27**	**88.85 ± 0.88**	**66.67 ± 1.59**	**68.40 ± 1.57**
HW	BRIDGE (ICCV'25)	79.74 ± 2.91	78.11 ± 0.54	71.37 ± 1.53	72.85 ± 2.04	70.92 ± 0.40	62.91 ± 1.27	64.83 ± 2.36	58.36 ± 0.46	48.20 ± 1.53	58.68 ± 1.72	50.17 ± 0.39	37.15 ± 1.82
	SCVT (TIP'25)	82.33 ± 0.53	74.98 ± 0.21	71.65 ± 0.29	80.97 ± 0.69	74.25 ± 0.26	70.64 ± 0.41	79.29 ± 0.68	73.46 ± 0.24	69.88 ± 0.30	77.61 ± 0.58	72.88 ± 0.26	69.25 ± 0.46
	FreeCSL (CVPR'25)	83.58 ± 1.81	85.76 ± 0.26	78.78 ± 0.56	80.87 ± 1.76	78.80 ± 0.26	70.61 ± 0.41	76.11 ± 1.39	72.74 ± 0.41	61.57 ± 0.50	72.12 ± 1.92	64.25 ± 0.39	54.94 ± 0.64
	GAVIM (TKDE'26)	88.53 ± 1.84	82.28 ± 1.39	78.55 ± 2.87	83.26 ± 2.60	73.33 ± 1.58	69.79 ± 2.70	77.41 ± 2.05	65.21 ± 2.15	61.11 ± 2.68	70.61 ± 1.71	59.93 ± 1.53	51.56 ± 2.54
	IMVC-NCD (ICML'26)	**95.15 ± 3.19**	89.55 ± 2.36	**89.65 ± 2.24**	86.06 ± 2.76	82.20 ± 1.39	78.54 ± 2.43	79.78 ± 3.20	72.77 ± 2.22	64.45 ± 2.38	71.19 ± 1.86	63.01 ± 1.97	53.43 ± 3.08
	DGIMVCM (AAAI'26)	94.97 ± 0.34	89.20 ± 0.31	89.26 ± 0.40	**89.95 ± 0.20**	84.58 ± 0.21	**85.08 ± 0.47**	**85.20 ± 0.22**	79.75 ± 0.27	77.18 ± 0.47	**82.29 ± 0.20**	**76.05 ± 0.21**	**72.77 ± 0.43**
	MVP (ICLR'25)	95.08 ± 0.47	**89.78 ± 0.36**	89.42 ± 0.73	86.77 ± 0.43	80.50 ± 0.24	81.75 ± 0.90	75.35 ± 0.31	72.48 ± 0.25	73.08 ± 0.72	68.75 ± 0.30	63.83 ± 0.31	65.41 ± 0.78
	CAMEL (ICLR'26)	92.33 ± 4.76	89.46 ± 0.86	89.46 ± 3.01	87.89 ± 4.68	**84.77 ± 0.92**	83.52 ± 2.18	84.61 ± 4.20	**79.86 ± 1.21**	**78.79 ± 2.57**	82.21 ± 4.36	73.24 ± 1.32	69.90 ± 2.20
	SACDG (Ours)	**95.46 ± 0.16**	**90.38 ± 0.44**	**90.12 ± 0.42**	**92.83 ± 0.13**	**85.91 ± 0.40**	**85.55 ± 0.45**	**89.77 ± 0.18**	**80.48 ± 0.49**	**79.25 ± 0.43**	**86.39 ± 0.18**	**76.37 ± 0.47**	**73.23 ± 0.30**
MNIST4	BRIDGE (ICCV'25)	86.31 ± 2.30	66.68 ± 1.53	68.17 ± 3.14	83.42 ± 2.15	64.66 ± 1.69	64.36 ± 2.70	80.40 ± 2.55	63.17 ± 1.60	61.26 ± 2.79	78.07 ± 2.71	61.13 ± 1.54	58.68 ± 3.08
	SCVT (TIP'25)	87.22 ± 0.30	70.56 ± 0.86	70.24 ± 0.84	85.49 ± 0.39	69.38 ± 0.69	69.31 ± 0.57	84.08 ± 0.35	68.36 ± 0.81	68.26 ± 0.60	83.42 ± 0.32	67.21 ± 0.68	67.57 ± 0.76
	FreeCSL (CVPR'25)	89.05 ± 0.96	73.59 ± 1.29	74.38 ± 1.21	88.58 ± 0.70	72.80 ± 1.01	73.61 ± 1.38	88.19 ± 0.58	71.75 ± 0.84	73.14 ± 1.98	87.88 ± 0.93	70.59 ± 0.85	71.96 ± 1.31
	GAVIM (TKDE'26)	89.64 ± 0.32	73.58 ± 1.05	75.11 ± 1.35	88.02 ± 0.26	69.29 ± 1.48	71.19 ± 1.01	85.85 ± 0.33	63.83 ± 1.43	67.11 ± 1.22	83.74 ± 0.28	60.58 ± 1.05	64.42 ± 1.52
	IMVC-NCD (ICML'26)	89.31 ± 0.14	73.52 ± 0.46	74.84 ± 0.38	89.14 ± 0.16	73.29 ± 0.60	74.52 ± 0.42	89.00 ± 0.17	73.02 ± 0.50	74.21 ± 0.47	88.85 ± 0.17	72.78 ± 0.50	73.99 ± 0.38
	DGIMVCM (AAAI'26)	90.83 ± 0.45	78.82 ± 0.50	78.36 ± 0.55	89.22 ± 0.39	75.05 ± 0.64	74.66 ± 0.37	85.77 ± 0.32	70.60 ± 0.75	72.05 ± 0.33	81.87 ± 0.34	67.55 ± 0.47	68.96 ± 0.43
	MVP (ICLR'25)	88.80 ± 0.20	71.44 ± 0.79	73.42 ± 0.18	87.07 ± 0.22	67.68 ± 1.03	68.01 ± 0.24	85.24 ± 0.25	64.04 ± 1.26	63.88 ± 0.24	82.46 ± 0.21	61.82 ± 0.94	59.14 ± 0.14
	CAMEL (ICLR'26)	**91.84 ± 0.47**	**80.10 ± 0.19**	**80.42 ± 0.11**	**91.31 ± 0.59**	**78.41 ± 0.13**	**79.50 ± 0.16**	**90.87 ± 0.57**	**76.44 ± 0.24**	**78.49 ± 0.12**	**90.51 ± 0.43**	**75.46 ± 0.15**	**77.36 ± 0.16**
	SACDG (Ours)	**92.31 ± 0.09**	**80.42 ± 0.17**	**81.14 ± 0.10**	**91.89 ± 0.08**	**79.11 ± 0.28**	**80.15 ± 0.17**	**91.44 ± 0.08**	**77.93 ± 0.21**	**79.07 ± 0.16**	**90.79 ± 0.08**	**76.86 ± 0.27**	**78.17 ± 0.11**
CUB	BRIDGE (ICCV'25)	54.42 ± 2.23	54.50 ± 3.81	41.65 ± 3.44	50.29 ± 1.64	50.23 ± 3.51	39.56 ± 2.93	45.59 ± 1.63	47.10 ± 3.56	35.80 ± 2.44	40.96 ± 2.00	41.40 ± 2.49	31.72 ± 2.16
	SCVT (TIP'25)	60.00 ± 1.41	62.88 ± 1.60	44.84 ± 1.51	53.41 ± 1.12	55.18 ± 1.58	39.65 ± 1.18	46.76 ± 1.65	46.54 ± 1.35	35.47 ± 1.06	38.49 ± 1.48	38.48 ± 1.42	29.71 ± 1.34
	FreeCSL (CVPR'25)	63.50 ± 1.06	60.72 ± 0.11	46.41 ± 0.75	57.23 ± 1.16	53.20 ± 0.14	38.42 ± 1.03	49.18 ± 1.31	47.65 ± 0.15	28.82 ± 0.62	40.75 ± 0.79	39.98 ± 0.13	22.18 ± 1.02
	GAVIM (TKDE'26)	75.00 ± 3.57	69.78 ± 2.27	58.41 ± 2.84	66.17 ± 2.55	63.98 ± 2.78	47.26 ± 3.50	54.92 ± 3.83	55.53 ± 3.57	35.64 ± 2.74	45.11 ± 2.66	45.82 ± 3.50	28.64 ± 2.67
	IMVC-NCD (ICML'26)	60.17 ± 2.35	57.50 ± 2.28	45.67 ± 2.80	48.48 ± 1.48	51.37 ± 3.08	38.28 ± 2.81	43.66 ± 2.18	44.97 ± 2.78	31.55 ± 2.06	37.78 ± 1.62	36.74 ± 2.85	24.43 ± 3.41
	DGIMVCM (AAAI'26)	71.67 ± 0.24	70.87 ± 0.28	62.16 ± 0.81	67.74 ± 0.23	66.33 ± 0.37	59.97 ± 1.03	62.72 ± 0.23	62.43 ± 0.37	58.11 ± 0.99	58.30 ± 0.27	58.27 ± 0.46	56.27 ± 1.05
	MVP (ICLR'25)	**82.08 ± 0.50**	77.08 ± 1.43	68.14 ± 1.33	74.73 ± 0.83	70.67 ± 1.83	**64.69 ± 1.98**	67.36 ± 0.58	67.07 ± 1.41	**62.51 ± 1.49**	61.91 ± 0.91	60.11 ± 2.30	**58.78 ± 2.07**
	CAMEL (ICLR'26)	81.25 ± 0.27	**81.88 ± 0.74**	**68.35 ± 0.92**	**77.36 ± 0.35**	**76.93 ± 0.69**	64.56 ± 0.88	**73.45 ± 0.21**	**72.91 ± 0.55**	62.03 ± 0.61	**68.83 ± 0.26**	**67.98 ± 0.55**	56.68 ± 0.93
	SACDG (Ours)	**82.50 ± 0.43**	**82.39 ± 0.49**	**69.48 ± 0.86**	**77.62 ± 0.54**	**77.35 ± 0.66**	**67.23 ± 0.53**	**73.70 ± 0.50**	**73.73 ± 0.43**	**62.80 ± 0.54**	**69.07 ± 0.53**	**68.20 ± 0.69**	**59.02 ± 0.82**
MNIST-Fea	BRIDGE (ICCV'25)	95.22 ± 1.14	92.42 ± 3.26	95.74 ± 1.49	92.18 ± 1.28	90.56 ± 3.04	93.27 ± 1.56	89.10 ± 0.86	88.52 ± 2.92	91.27 ± 1.92	84.39 ± 0.99	86.38 ± 3.19	88.08 ± 1.17
	SCVT (TIP'25)	97.42 ± 0.29	93.87 ± 0.31	96.16 ± 1.08	96.56 ± 0.31	93.09 ± 0.39	93.65 ± 0.72	95.01 ± 0.36	92.32 ± 0.40	90.97 ± 0.85	93.48 ± 0.32	91.36 ± 0.43	87.45 ± 0.92
	FreeCSL (CVPR'25)	98.13 ± 0.23	96.33 ± 0.46	97.09 ± 0.18	89.92 ± 0.21	90.87 ± 0.56	89.56 ± 0.11	85.43 ± 0.15	85.50 ± 0.46	80.85 ± 0.16	80.62 ± 0.15	78.78 ± 0.53	74.48 ± 0.10
	GAVIM (TKDE'26)	98.27 ± 0.05	95.10 ± 0.13	96.20 ± 0.09	97.93 ± 0.04	94.60 ± 0.17	96.01 ± 0.10	97.56 ± 0.06	94.08 ± 0.10	95.62 ± 0.13	97.38 ± 0.04	93.62 ± 0.11	95.04 ± 0.12
	IMVC-NCD (ICML'26)	**98.75 ± 0.04**	**96.37 ± 0.11**	**97.26 ± 0.08**	**98.23 ± 0.04**	**95.72 ± 0.12**	**96.60 ± 0.09**	**97.97 ± 0.05**	**94.81 ± 0.09**	**96.06 ± 0.08**	**97.47 ± 0.04**	**94.29 ± 0.11**	**95.05 ± 0.13**
	DGIMVCM (AAAI'26)	98.13 ± 0.39	95.10 ± 0.51	96.20 ± 0.39	96.74 ± 0.61	94.46 ± 0.51	95.48 ± 0.57	95.31 ± 0.47	94.05 ± 0.53	94.86 ± 0.57	93.68 ± 0.61	93.51 ± 0.41	94.38 ± 0.45
	MVP (ICLR'25)	98.29 ± 0.05	95.26 ± 0.11	96.30 ± 0.05	98.06 ± 0.05	94.85 ± 0.11	95.97 ± 0.06	97.79 ± 0.04	94.36 ± 0.12	95.55 ± 0.08	97.38 ± 0.04	94.12 ± 0.12	95.05 ± 0.06
	CAMEL (ICLR'26)	84.75 ± 0.23	66.98 ± 0.27	66.52 ± 0.30	78.97 ± 0.22	66.50 ± 0.34	66.38 ± 0.38	73.18 ± 0.24	66.08 ± 0.34	66.28 ± 0.28	68.64 ± 0.27	65.76 ± 0.28	66.08 ± 0.30
	SACDG (Ours)	**99.15 ± 0.04**	**96.89 ± 0.04**	**97.66 ± 0.03**	**98.64 ± 0.05**	**96.20 ± 0.04**	**97.08 ± 0.03**	**98.32 ± 0.04**	**95.34 ± 0.05**	**96.35 ± 0.04**	**97.81 ± 0.04**	**94.80 ± 0.05**	**95.52 ± 0.05**
Scene15	BRIDGE (ICCV'25)	34.42 ± 0.38	36.55 ± 0.20	19.44 ± 0.92	32.81 ± 0.36	35.31 ± 0.28	18.24 ± 0.91	30.88 ± 0.56	34.11 ± 0.24	17.23 ± 0.98	29.81 ± 0.58	32.87 ± 0.23	16.20 ± 0.86
	SCVT (TIP'25)	38.35 ± 1.69	41.28 ± 1.58	21.89 ± 1.63	37.19 ± 2.05	39.71 ± 1.81	20.75 ± 1.48	36.08 ± 1.97	38.01 ± 1.28	19.73 ± 1.58	34.83 ± 1.67	36.52 ± 1.17	18.45 ± 1.11
	FreeCSL (CVPR'25)	36.08 ± 0.89	38.97 ± 0.69	21.68 ± 1.04	33.31 ± 0.87	36.62 ± 0.81	19.33 ± 0.97	31.77 ± 0.73	33.64 ± 0.64	17.36 ± 0.88	29.11 ± 0.76	32.06 ± 0.76	15.08 ± 0.79
	GAVIM (TKDE'26)	39.81 ± 1.20	38.73 ± 1.12	23.31 ± 1.51	38.28 ± 0.91	37.61 ± 0.96	22.33 ± 1.34	37.06 ± 1.00	36.07 ± 1.74	21.41 ± 1.58	35.53 ± 1.09	34.81 ± 1.42	20.17 ± 1.65
	IMVC-NCD (ICML'26)	44.71 ± 0.94	46.77 ± 0.36	29.04 ± 0.45	43.17 ± 1.29	45.15 ± 0.32	27.97 ± 0.71	42.29 ± 1.55	**44.30 ± 0.45**	**26.68 ± 0.77**	41.46 ± 1.35	**42.85 ± 0.53**	**25.70 ± 0.67**
	DGIMVCM (AAAI'26)	45.48 ± 1.77	45.76 ± 0.91	28.50 ± 0.87	41.65 ± 1.56	41.38 ± 0.81	24.23 ± 0.76	36.83 ± 1.57	38.70 ± 0.64	20.82 ± 0.88	33.19 ± 1.18	36.45 ± 0.83	18.58 ± 1.21
	MVP (ICLR'25)	43.20 ± 1.52	44.26 ± 0.19	26.32 ± 0.70	38.85 ± 1.77	40.74 ± 0.17	24.38 ± 0.65	37.07 ± 1.71	37.91 ± 0.17	22.29 ± 0.50	34.65 ± 1.89	36.12 ± 0.19	18.75 ± 0.67
	CAMEL (ICLR'26)	**51.68 ± 0.45**	**49.09 ± 0.33**	**31.97 ± 0.28**	**49.46 ± 0.34**	**46.58 ± 0.46**	**29.15 ± 0.40**	**47.63 ± 0.41**	44.24 ± 0.47	25.93 ± 0.25	**44.22 ± 0.29**	41.70 ± 0.37	23.16 ± 0.39
	SACDG (Ours)	**52.09 ± 0.36**	**49.54 ± 0.71**	**32.74 ± 0.52**	**50.28 ± 0.41**	**47.38 ± 1.08**	**30.20 ± 0.36**	**48.16 ± 0.42**	**45.57 ± 0.89**	**28.34 ± 0.40**	**45.20 ± 0.38**	**43.62 ± 0.90**	**26.22 ± 0.40**
Reuters	BRIDGE (ICCV'25)	50.27 ± 1.29	28.52 ± 0.04	18.76 ± 0.04	48.06 ± 1.35	27.89 ± 0.04	17.87 ± 0.04	45.60 ± 1.42	27.06 ± 0.04	17.14 ± 0.04	42.65 ± 1.46	26.73 ± 0.05	16.49 ± 0.03
	SCVT (TIP'25)	50.27 ± 0.37	29.33 ± 0.62	22.65 ± 0.62	48.10 ± 0.50	28.66 ± 0.66	21.59 ± 0.53	46.01 ± 0.52	28.03 ± 0.45	20.81 ± 0.51	43.28 ± 0.60	27.37 ± 0.55	19.81 ± 0.42
	FreeCSL (CVPR'25)	49.17 ± 0.27	27.01 ± 0.37	17.09 ± 0.30	45.42 ± 0.25	26.61 ± 0.31	16.22 ± 0.33	43.07 ± 0.25	26.14 ± 0.26	15.56 ± 0.38	39.83 ± 0.28	25.76 ± 0.36	14.62 ± 0.38
	GAVIM (TKDE'26)	50.25 ± 0.19	29.33 ± 0.44	22.65 ± 0.35	48.70 ± 0.16	28.82 ± 0.37	21.59 ± 0.43	46.05 ± 0.19	27.83 ± 0.52	20.66 ± 0.41	42.66 ± 0.18	26.76 ± 0.39	19.71 ± 0.42
	IMVC-NCD (ICML'26)	50.27 ± 0.62	29.33 ± 0.54	22.65 ± 0.57	47.85 ± 0.41	28.83 ± 0.59	21.50 ± 0.55	45.07 ± 0.56	28.13 ± 0.49	20.51 ± 0.44	42.18 ± 0.46	27.60 ± 0.58	19.31 ± 0.46
	DGIMVCM (AAAI'26)	50.27 ± 0.39	29.33 ± 0.38	22.65 ± 0.46	47.18 ± 0.53	28.78 ± 0.61	21.86 ± 0.51	45.13 ± 0.63	28.29 ± 0.59	21.02 ± 0.64	41.68 ± 0.63	27.58 ± 0.50	19.96 ± 0.43
	MVP (ICLR'25)	51.66 ± 1.99	30.07 ± 2.36	23.69 ± 3.09	50.23 ± 2.25	29.79 ± 2.36	22.50 ± 2.07	48.96 ± 1.79	29.47 ± 2.05	21.84 ± 2.92	**48.14 ± 2.39**	**29.27 ± 2.15**	20.91 ± 2.18
	CAMEL (ICLR'26)	**55.77 ± 0.42**	**35.63 ± 0.16**	**31.74 ± 0.36**	**52.31 ± 0.42**	**33.39 ± 0.21**	**29.77 ± 0.31**	**50.00 ± 0.57**	**31.00 ± 0.23**	**28.23 ± 0.32**	46.35 ± 0.62	28.65 ± 0.25	**25.65 ± 0.27**
	SACDG (Ours)	**61.83 ± 0.54**	**38.51 ± 0.54**	**35.27 ± 0.79**	**59.54 ± 0.78**	**36.66 ± 0.43**	**33.64 ± 0.64**	**57.63 ± 0.66**	**34.66 ± 0.38**	**31.81 ± 0.77**	**55.47 ± 0.55**	**32.80 ± 0.33**	**30.61 ± 0.64**

Table 2: Ablation results (mean ± std, %) on MNIST-Fea, Scene15, and Reuters at a 10% missing rate. Best and second-best results are highlighted in red and blue, respectively.

Losses			MNIST-Fea			Scene15			Reuters		
L_{rec}	L_{imp}	L_{FCA}	ACC	NMI	ARI	ACC	NMI	ARI	ACC	NMI	ARI
✓	×	×	**91.27 ± 4.21**	**87.96 ± 1.65**	**90.30 ± 4.19**	31.86 ± 0.96	32.08 ± 1.53	24.56 ± 1.49	36.51 ± 3.28	16.81 ± 0.30	29.58 ± 0.77
✓	×	✓	89.81 ± 3.04	85.98 ± 0.46	86.84 ± 3.00	**32.02 ± 0.29**	32.45 ± 0.82	**25.21 ± 0.47**	**38.72 ± 1.02**	**21.10 ± 1.57**	**30.89 ± 0.66**
✓	✓	×	89.42 ± 5.84	85.99 ± 3.09	87.44 ± 5.82	31.77 ± 1.11	**33.19 ± 0.91**	24.79 ± 1.73	29.30 ± 0.62	8.52 ± 1.44	22.73 ± 0.21
✓	✓	✓	**99.15 ± 0.04**	**96.89 ± 0.04**	**97.66 ± 0.03**	**52.09 ± 0.36**	**49.54 ± 0.71**	**32.74 ± 0.52**	**61.83 ± 0.54**	**38.51 ± 0.54**	**35.27 ± 0.79**

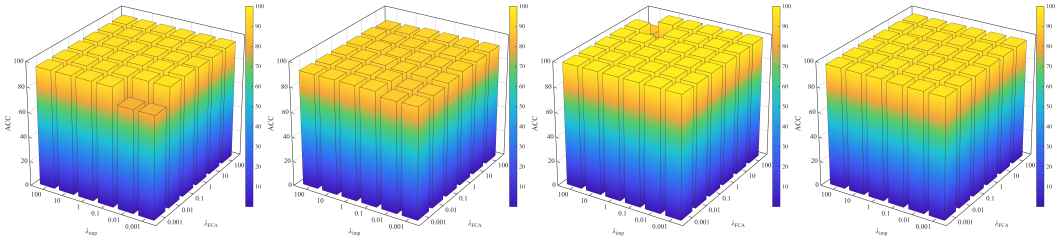
Reuters, respectively. The consistent advantage of the full objective across the three metrics shows that both missing view completion and the alignment of fusion and clustering contribute to the final clustering performance. This further confirms the complementary roles of both objectives during optimization and highlights their joint importance for stable clustering under incompleteness.

3) Parameter Analysis: Fig. 2 shows broad regions of high ACC with respect to λ_{imp} and λ_{FCA}, indicating that SACDG remains stable under moderate variations in the weights for completion, fusion, and clustering. These results demonstrate the robustness of the proposed framework across different parameter configurations without requiring precise tuning of either coefficient.

4) Convergence Analysis: As shown in Fig. 3, SACDG improves rapidly and reaches stable performance on the four representative datasets under a 10% missing rate. This verifies the effectiveness of the joint optimization strategy in achieving stable representation learning and clustering.

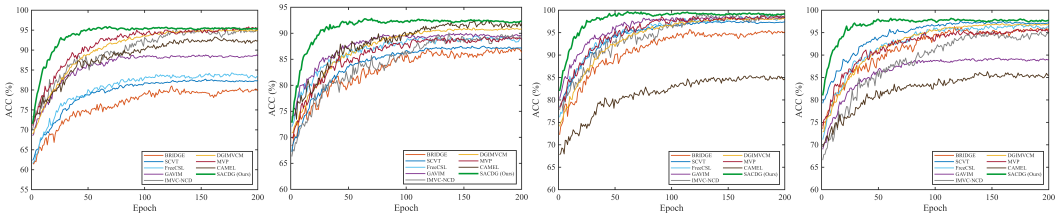
5) Representation Visualization: Fig. 4 shows that the MNIST4 representations become progressively more compact and separable during training, with ACC increasing from 56.35% to 92.50%. These results indicate that SACDG gradually enhances the discriminative capability of learned representations while preserving meaningful semantic structures across samples throughout the training.

6) t-SNE Comparison: Fig. 5 further shows clearer and more compact MNIST-Fea clusters than the compared methods, consistent with the quantitative results in Table 1. This suggests that the recovered representations preserve more discriminative cluster structures under incomplete observations. Moreover the improved visualization reflects that SACDG effectively captures latent semantics and maintains consistent inter-cluster separation across incomplete multi-view data distributions.



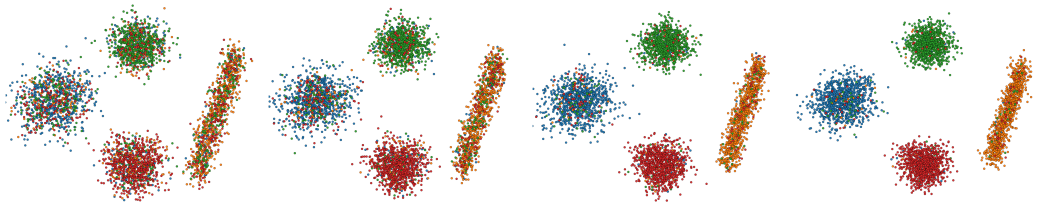
(a) HW (b) MNIST4 (c) MNIST-Fea (d) Synthetic3D

Figure 2: Parameter sensitivity of our method on four datasets.



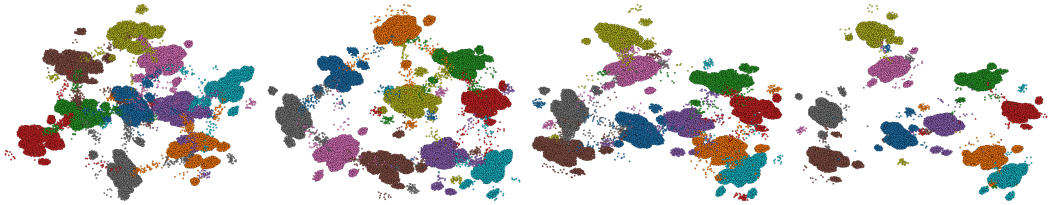
(a) HW (b) MNIST4 (c) MNIST-Fea (d) Synthetic3D

Figure 3: Convergence comparison of different methods on four datasets under a 10% missing rate.



(a) 25 epoch (56.35%) (b) 50 epoch (70.28%) (c) 75 epoch (86.18%) (d) 100 epoch (92.50%)

Figure 4: t-SNE visualization of the learned representations on MNIST4 at different training epochs.



(a) CAMEL (b) GAVIM (c) IMVC-NCD (d) SACDG (Ours)

Figure 5: t-SNE visualization on MNIST-Fea at a 10% missing rate.

5 Conclusion

In this paper, we presented SACDG, a unified framework for incomplete multi-view clustering. Rather than reconstructing unavailable views, SACDG models recovery as conditional drift from a cross-view semantic anchor toward the target-view distribution, guided by local graph context and prototype semantics. This design preserves sample correspondence while recovering target-specific structure. Reliability estimation calibrates recovered representations to reduce the influence of uncertain completions during fusion. Graph alignment and prototype-driven clustering provide structural and semantic feedback to the generator, coupling recovery and clustering within an iterative process. Experiments on seven benchmarks under multiple missing rates demonstrate the effectiveness and robustness of SACDG under severe incompleteness. Future work will explore larger-scale and heterogeneous settings while improving computational efficiency and uncertainty calibration.

REFERENCES

Hongmin Cai, Weitian Huang, Sirui Yang, Siqi Ding, Yue Zhang, Bin Hu, Fa Zhang, and Yiu-Ming Cheung. Realize generative yet complete latent representation for incomplete multi-view learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(5):3637–3652, 2023.

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*, pp. 11221–11229, 2024.

Jie Chen, Hua Mao, Wai Lok Woo, and Xi Peng. Deep multiview clustering by contrasting cluster assignments. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 16706–16715, 2023.

Wenlan Chen, Lu Gao, Daoyuan Wang, Fei Guo, and Cheng Liang. Geometry-aware variational information maximization for deep incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 20289–20297, 2026.

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*, pp. 5071–5081, 2025.

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, 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*, pp. 20941–20949, 2026.

Ran Eisenberg, Jonathan Svirsky, and Ofir Lindenbaum. Coper: Correlation-based permutations for multi-view clustering. In *International Conference on Learning Representations*, pp. 46154–46179, 2025.

Xin Gao and Jian Pu. Deep incomplete multi-view learning via cyclic permutation of vaes. In *International Conference on Learning Representations*, pp. 99609–99643, 2025.

Zhibin Gu, Zhenhao Zhong, Xi Zhang, and Bing Li. Constant degree matrix-driven incomplete multi-view clustering via connectivity-structure and embedding tensor learning. In *International Conference on Learning Representations*, pp. 84649–84671, 2026.

Qian Guo, Gaohui Zuo, Bingbing Jiang, Guangrui Fan, Zhihua Cui, Xinyan Liang, and Jianjian Ding. Incomplete multi-view clustering via neighborhood-conditioned diffusion. In *International Conference on Machine Learning*, 2026.

Bingbing Jiang, Chenglong Zhang, Xinyan Liang, Peng Zhou, Jie Yang, Xingyu Wu, Junyi Guan, Weiping Ding, and Weiguo Sheng. Collaborative similarity fusion and consistency recovery for incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 17617–17625, 2025a.

Xiaorui Jiang, Buyun He, Peng Yuan Zhou, Xinyue Chen, Jingcai Guo, Jie Xu, and Yong Liao. A unified framework to bridge complete and incomplete deep multi-view clustering under non-iid missing patterns. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 594–603, 2025b.

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*, pp. 1016–1026, 2025.

X. Li, G. Li, X. Zhang, J. Wen, and L. Wu. Crossnet-vga: Variational collaboration and graph attention fusion for incomplete multi-view clustering. *IEEE Transactions on Image Processing*, 34:7448–7461, 2025.

Zhen Long, Qiyuan Wang, Yazhou Ren, Yipeng Liu, and Ce Zhu. S2mvtc: A simple yet efficient scalable multi-view tensor clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 24088–24097, 2024.

Jingyu Pu, Chenhang Cui, Xinyue Chen, Yazhou Ren, Xiaorong Pu, Zhifeng Hao, Philip S Yu, and Lifang He. Adaptive feature imputation with latent graph for deep incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 14633–14641, 2024.

Yalan Qin, Guorui Feng, and Xinpeng Zhang. Scalable one-pass incomplete multi-view clustering by aligning anchors. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 20042–20050, 2025.

Yuze Tan, Yixi Liu, Shudong Huang, Wentao Feng, and Jiancheng Lv. Sample-level multi-view graph clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 23966–23975, 2023.

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

Bocheng Wang, Chusheng Zeng, Mulin Chen, and Xuelong Li. Towards learnable anchor for deep multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 21044–21052, 2025.

Qianqian Wang, Xurui Liao, Wei Feng, and Quanxue Gao. Debiased and denoised representation learning for incomplete multi-view clustering. In *International Conference on Learning Representations*, pp. 73338–73350, 2026.

Yuanzhi Wang, Zhen Cui, and Yong Li. Distribution-consistent modal recovering for incomplete multimodal learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 21968–21977, 2023.

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*, 2024.

Tingting Wu, Zhendong Li, Zhibin Gu, Jiazheng Yuan, and Songhe Feng. Koala: Kernel coupling and element imputation induced multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 21581–21589, 2025.

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*, 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*, 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*, pp. 8761–8769, 2022.

Weiqing Yan, Yuanyang Zhang, Chenlei Lv, Chang Tang, Guanghui Yue, Liang Liao, and Weisi Lin. Gcfagg: Global and cross-view feature aggregation for multi-view clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 19863–19872, 2023.

Xiaoqiang Yan, Zhixiang Jin, Fengshou Han, and Yangdong Ye. Differentiable information bottleneck for deterministic multi-view clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 27425–27434, 2024.

Chao Zhang, Huaxiong Li, Wei Lv, Zizheng Huang, Yang Gao, and Chunlin Chen. Enhanced tensor low-rank and sparse representation recovery for incomplete multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 11174–11182, 2023.

Teng Zhang, Pengyuan Li, Zisen Kong, Dongxia Chang, and Yao Zhao. Bipartite contrastive multi-view clustering with singular value modulation. *Neurocomputing*, 656:131528, 2025a.

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*, pp. 22650–22658, 2025b.

Yunhe Zhang, Jinyu Cai, Zhihao Wu, Pengyang Wang, and See-Kiong Ng. Mixture of experts as representation learner for deep multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 22704–22713, 2025c.

Zhenghao Zhang, Jun Xie, Xingchen Chen, Tao Yu, Hongzhu Yi, Kaixin Xu, Yuanxiang Wang, Tianyu Zong, Xinming Wang, Jiahuan Chen, et al. Dynamic deep graph learning for incomplete multi-view clustering with masked graph reconstruction loss. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 28600–28608, 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*, pp. 11327–11335, 2023.

Yunxiao Zhao, Liang Bai, and Xian Yang. Adaptive evolutionary fusion for multi-view clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 28769–28777, 2026.