Uncertainty-Aware Graph Contrastive Clustering with Coupled Dual-Path Augmentation

Yanpei Xiao[1], Shangshang Zhao[1], Feng Liu[1], Zhongyang Zhou[1], and Feiyu Chen[2,*]
[1]College of Computer and Infomation Science, Chongqing Normal University, Chongqing, 400030, China
[2]National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing, 400030, China
Email: { yanpei_xiao, zs1999ism, lf_ism, zhouzhongyoung, fchen_cqnu}@163.com

***Abstract*—Graph contrastive clustering has emerged as a promising paradigm for unsupervised graph representation learning, which leverages contrastive objectives to learn discriminative embeddings for high-quality clustering. Despite its effectiveness, most existing methods rely on predefined graph augmentation strategies and do not explicitly take the downstream clustering task into account. In practical scenarios, this inevitably introduces noise or distorts the underlying semantics. Such inherent limitations often lead to semantic drift between augmented graph views, ultimately degrading clustering performance. To address these issues, we propose a Uncertainty-Aware Graph Contrastive Clustering with Coupled Dual-Path Augmentation method, termed UGCC. The proposed method constructs multi-view contrastive learning objectives through learnable attribute augmentation and adaptive structure augmentation, enabling the generation of high-quality augmented samples and the learning of more discriminative representations for graph contrastive clustering. Furthermore, we introduce an uncertainty-aware clustering alignment objective to adaptively regulate the optimization process of unsupervised clustering. Extensive experiments conducted on five benchmark datasets demonstrate the effectiveness of UGCC.**

***Index Terms*—Contrastive Learning, Graph Neural Network, Deep Graph Clustering**

I. Introduction

Graph-structured data are ubiquitous in the real world, ranging from social and recommender systems to biomolecular analysis, highlighting their practical importance [1]. Given that a large amount of real-world graph data lack labeled information, how to effectively capture complex dependencies among nodes under unsupervised settings and learn node representations that preserve structural information while maintaining semantic discriminability for high-performance graph clustering has become a key challenge in graph learning.

A representative line of research is deep graph clustering, which integrates deep representation learning with clustering objectives. Early methods mainly adopt autoencoder-based reconstruction to fuse attribute and topological information. For example, SDCN [2] combines a deep autoencoder with a GCN and introduces self-supervised consistency for graph clustering. More recent approaches focus on improving representation quality and optimization stability. AGC-DRR [3] reduces redundancy in latent and intermediate spaces and enforces clustering consistency. Despite these advances, deep graph clustering remains challenged by noisy graph structures, sparse features, and unreliable clustering supervision in fully unsupervised settings, where pseudo-labels or intermediate assignments can be unstable and error-amplifying in early training stages [4]. In parallel, Graph Contrastive Learning (GCL) has emerged as an effective self-supervised paradigm for unsupervised representation learning by contrasting different views [5]. Classic methods such as MVGRL [6] learn node/graph representations by maximizing agreement across views. Recent studies further show that performance is highly dependent on the quality and task-relevance of constructed views: NCLA [7] leverages neighborhood-level objectives with learnable augmentation, while HSAN [8] exploits higher-order similarity with an augmentation-free contrastive scheme. Building on these advances, graph contrastive clustering [9] integrates contrastive learning with clustering to obtain cluster-friendly embeddings. For example, SECL [10] constructs attribute/structure perspectives and injects community information via modularity maximization, HomoCAGC [11] exploits homophily to refine positive neighbor sets under noisy edges, and AutoSSL [12] introduces adaptive augmentations with cross-view self-labeling to reduce reliance on handcrafted augmentations.

Nevertheless, existing graph contrastive clustering methods [13]–[15] still heavily depend on contrastive sample quality. Many approaches use heuristic, task-agnostic augmentations (e.g., feature perturbation, edge removal, and random sampling) that are not adapted to the model state or clustering needs; we refer to them as non-adaptive augmentation (NAA). With fixed forms and intensities, NAA may distort semantics, disrupt topology, or introduce noise, and the resulting views may be suboptimal for clustering. Moreover, fully unsupervised clustering often relies on pseudo-labels [4] or intermediate assignments, which can be unreliable early in training and amplify errors. Consequently, generating clustering-relevant augmented samples and robustly optimizing unsupervised clustering objectives remain key challenges.

To address the aforementioned limitations, we propose Uncertainty-Aware Graph Contrastive Clustering with Coupled Dual-Path Augmentation (UGCC). UGCC integrates learnable attribute augmentation and adaptive structure augmentation, avoiding reliance on predefined schemes. In addition, we introduce a mutually balanced objective to guarantee the consistency of the embeddings as well as the diversity of

This work was supported by the Natural Science Foundation of Chongqing (Grant No. CSTB2025NSCQ-LZX0068,CSTB2025NSCQ-LZX0056).
* Corresponding Author

the augmented samples. Finally, we design an uncertainty-aware clustering alignment objective to stabilize unsupervised clustering under noisy pseudo-labels. This objective adaptively weights clustering alignment constraints based on prediction uncertainty, assigning higher weights to more confident samples and down-weighting uncertain ones to improve training stability.

The main contributions of this paper are:

- We propose UGCC, an efficient deep graph contrastive clustering method that performs adaptive multi-view augmentation and clustering-oriented optimization in an end-to-end manner, without pre-training.
- We introduce a coupled dual-path augmentation module (CDPA) that simultaneously conducts learnable attribute augmentation and adaptive structure augmentation. By jointly constructing multi-view contrastive learning objectives, this module enhances the discriminability of learned representations.
- We design an uncertainty-aware clustering alignment strategy to mitigate the instability caused by noisy pseudo-labels and stabilize unsupervised clustering optimization without discarding training data.

II. METHODOLOGY

The overall framework of UGCC is shown in Fig 1. The main components of the proposed method include the coupled dual-path augmentation module and the uncertainty-aware clustering alignment strategy.

A. Coupled Dual-Path Augmentation Module

In graph contrastive clustering, effective view construction requires sufficient diversity while preserving clustering semantics. However, most existing methods rely on perturbation-based augmentation (PBA), which mainly depends on random perturbations or heuristic rules and often overlooks the complementary roles of attribute semantics and graph topology. Consequently, the resulting view discrepancies may be misaligned with clustering objectives, weakening contrastive learning. To address this issue, we propose Coupled Dual-Path Augmentation (CDPA), which jointly generates learnable augmented feature views in the attribute space and adaptively learned augmented topology views in the structural space. By jointly optimizing both pathways, CDPA produces multi-view differences that are better aligned with clustering semantics.

a) ***Structure Augmentation****:* Before encoding, we utilize the original attribute matrix X to compute a k-nearest-neighbor adjacency matrix A_{knn} using the kNN algorithm, so as to capture potential topological relationships among nodes in the original feature space. Meanwhile, to adaptively fuse the original graph structure with the kNN-based structural information, we introduce an attention mechanism at the adjacency-matrix level. Specifically, the original adjacency matrix is first processed by a self-attention mechanism to obtain an intermediate matrix A_{self}. Then, A_{self} and A_{knn} are further integrated through a cross-attention mechanism to produce the fused adjacency matrix A_{cross} . The detailed formulation is given as follows:

$$A_{self} = ReLU\left(\frac{(AW_q)(AW_k)^\top}{\sqrt{h}}\right), \tag{1}$$

where W_q and W_k denote learnable projection matrices, and hrepresents the hidden dimensionality of the attention space.

Subsequently, we introduce a cross-attention mechanism, where the original adjacency matrix A serves as the query and the kNN graph A_{knn} acts as the key. This design selectively incorporates structure relations induced by feature similarity while preserving the original topological semantics.

$$A_{cross} = ReLU\left(\frac{(AW_q)(A_{knn}W_k)^\top}{\sqrt{h}}\right). \tag{2}$$

After obtaining the self-attention matrix and the cross-attention matrix, we concatenate them and fuse the concatenated representation through a multilayer perceptron (MLP) to generate a joint representation:

$$A_{joint} = MLP([A_{self} \,||\, A_{cross}]) \in \mathbb{R}^{N\times N}, \tag{3}$$

where $\|$ denotes the concatenation operation.

Considering that the original graph structure still contains important prior information, we use a residual connection to fuse the augmented structure with the original structure:

$$A' = \gamma A_{joint} + (1-\gamma)A, \tag{4}$$

where $\gamma \in [0,1]$ denotes the residual coefficient, which is used to balance the contributions of the augmented structure and the original structure. This design introduces complementary structural information while effectively preventing excessive distortion of the original topological relationships.

The resulting matrix A' is used as the augmented adjacency matrix for subsequent graph representation learning. Note that the attention mechanism is applied at the adjacency-matrix level to model structural relationships, and therefore no explicit value matrix V is introduced. Through this structural augmentation module, the model incorporates latent feature-induced structural relations while preserving the original graph topology, providing a more stable foundation for representation learning and clustering.

b) ***Attribute Augmentation****:* To construct attribute-level augmented views, we introduce a learnable attribute augmentation mechanism that generates enhanced feature representations through a parameterized mapping of node attributes, providing flexible and task-relevant views for contrastive learning. Unlike PBA, this mechanism is integrated into model optimization and can be adaptively adjusted during training to better support clustering objectives. Specifically, given the original node attribute matrix $X \in \mathbb{R}^{N\times d}$, the augmented attributes are generated via a mapping function $f_\theta(\cdot)$ parameterized by an MLP:

$$X' = f_\theta(X), \tag{5}$$

where θ denotes the learnable parameters of the attribute augmentation module. The attribute augmentation module is

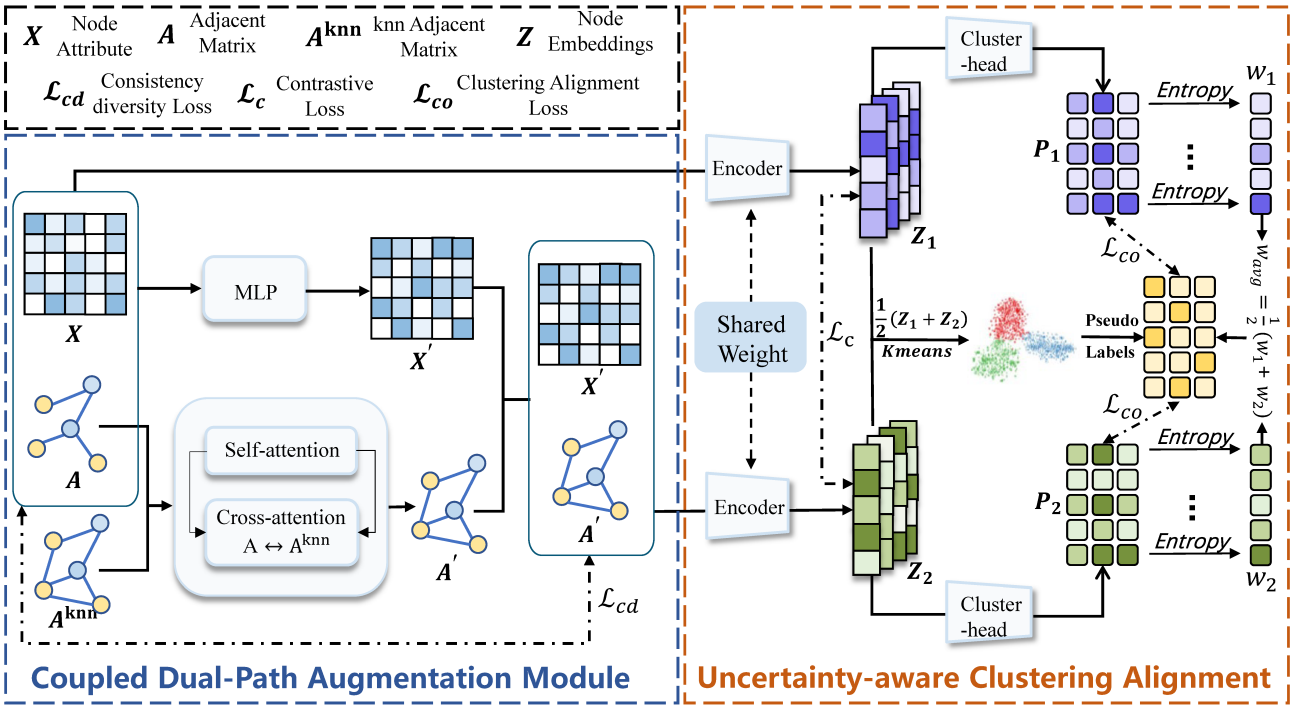


Fig. 1. UGCC constructs two complementary views via coupled dual-path augmentation in the attribute and structural spaces, followed by shared encoding and contrastive learning. An uncertainty-aware clustering alignment strategy is further applied to stabilize unsupervised clustering and obtain clustering-oriented representations.

jointly optimized with the encoder in an end-to-end manner. Gradients from the contrastive loss are backpropagated to θ, enabling the augmentation mechanism to adapt to representation learning.

B. Uncertainty-aware Clustering Alignment

Although clustering alignment can guide representation learning, noisy cluster assignments in the early stages of training may undermine its effectiveness. Directly enforcing clustering consistency on all samples can amplify erroneous alignments and destabilize the optimization process. To address this issue, we introduce an uncertainty-aware clustering alignment objective. Built upon the contrastive learning framework, this objective adaptively weights clustering alignment constraints according to the uncertainty of the model's predictive distribution. Samples with lower uncertainty are assigned higher weights, while those with higher uncertainty are down-weighted, thereby suppressing the influence of unreliable signals and improving training stability.

a) ***Pseudo-label generation from embeddings****:* To obtain pseudo-labels, we do not rely on the output of a single view, but instead leverage the consensus information from two views. Specifically, we first use the encoder network to obtain the embedding representations of the original graph $G = (X, A)$ and the augmented graph $G' = (X', A')$, which are computed as follows:

$$Z^{(1)} = \mathcal{F}(G), \quad Z^{(2)} = \mathcal{F}(G'). \tag{6}$$

Then, we fuse the embedding representations from the two views as follows:

$$Z = \frac{1}{2}(Z^{(1)} + Z^{(2)}), \tag{7}$$

based on Z, we apply the K-means clustering algorithm to obtain a pseudo-label $\widehat{y}_i \in [1, 2, ..., K]$ for each node i, where K denotes the predefined number of clusters.

b) ***Prediction logit computation****:* Given the embedding representations $Z^{(1)}$ and $Z^{(2)}$ obtained from the encoder, we project them into the clustering space using a multilayer perceptron (MLP), which can be expressed as follows:

$$P^{(v)} = Softmax\left(Z^{(v)} W_c^{\top}\right), \quad v \in \{1, 2\}, \tag{8}$$

where $Z^{(v)} \in \mathbb{R}^{N \times d}$, $W_c \in \mathbb{R}^{K \times d}$. $P^{(v)} \in \mathbb{R}^{N \times K}$ and $p_{i,k}^{(v)}$ denotes the probability of assigning node i to cluster k under view v.

c) ***Entropy-based uncertainty estimation****:* According to information theory, the entropy of the predicted distribution provides an effective measure of model uncertainty. For node i, the entropy under view v is defined as:

$$w_i^{(v)} = 1 - \frac{H\left(p_i^{(v)}\right)}{\log K} = 1 + \frac{1}{\log K} \sum_{k=1}^{K} p_{i,k}^{(v)} \log p_{i,k}^{(v)}, \quad v \in \{1, 2\}, \tag{9}$$

where $p_i^{(v)} = [p_{i,1}^{(v)}, \ldots, p_{i,K}^{(v)}]$ denotes the predicted cluster distribution of node i under view v, and $\log K$ is the maximum possible entropy. Thus, $w_i^{(v)} \in [0, 1]$, where a larger value indicates higher confidence or lower uncertainty.

To further improve stability, we use the average confidence across the two views as the final sample weight $w_i = \frac{1}{2}\left(w_i^{(1)} + w_i^{(2)}\right)$. This design avoids misjudgments caused by incidental perturbations in a single view. It is worth emphasizing that, unlike confidence-threshold-based sample selection methods, our approach does not discard any samples. Instead, the influence of each sample is smoothly modulated through continuous weighting. This uncertainty-aware weighting mechanism effectively suppresses the negative impact of potentially noisy pseudo-labels while preserving the continuity of the optimization process.

*d) **Uncertainty-aware clustering alignment**:* Finally, we incorporate the computed uncertainty weights into the cross-entropy loss. The objective is to minimize the weighted clustering prediction error, encouraging the model to prioritize confident samples while automatically down-weighting the influence of ambiguous ones. The final uncertainty-aware clustering alignment loss is defined as:

$$\mathcal{L}_{co} = \frac{1}{2N}\sum_{i=1}^{N} w_i\Big(CE(\widehat{y}_i,\ p_i^{(1)}) + CE(\widehat{y}_i,\ p_i^{(2)})\Big). \quad (10)$$

This objective encourages the model to prioritize clustering consistency learning for samples with high prediction confidence, while automatically reducing the constraint strength for uncertain samples. As a result, it helps stabilize the optimization of clustering structures under unsupervised settings.

C. Optimization

The proposed UGCC follows the general contrastive learning paradigm, aiming to maximize consistency across different views. Specifically, the overall objective consists of three loss functions: a consistency–diversity balancing loss $\mathcal{L}_{cd}$, a contrastive loss $\mathcal{L}_c$, and an uncertainty-aware clustering alignment loss $\mathcal{L}_{co}$.

To prevent representation collapse caused by overly strong consistency constraints in contrastive learning, we introduce a consistency–diversity balancing loss $\mathcal{L}_{cd}$ into the contrastive objective. This loss plays an adversarial push-pull regularization role: while the contrastive objective encourages cross-view invariance, $\mathcal{L}_{cd}$ encourages the augmented view to remain sufficiently different from the original graph, thereby avoiding trivial view generation and maintaining informative contrastive signals. Specifically, $\mathcal{L}_{cd}$ is defined as the negative mean squared error between the original graph G and the augmented graph G', which can be formulated as:

$$\mathcal{L}_{cd} = -(\|X' - X\|_F^2 + \|A' - A\|_F^2), \quad (11)$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

In this method, the contrastive learning backbone adopts a temperature-scaled cross-view contrastive loss. For each node i, a positive pair $(Z_i^{(1)}, Z_i^{(2)})$ is constructed across two views, while the representations of all other nodes are treated as negative samples. The contrastive loss can be defined as:

$$\mathcal{L}_c = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{\exp\Big(\mathrm{sim}\Big(Z_i^{(1)}, Z_i^{(2)}\Big)/T\Big)}{\sum_{k=1,\,k\neq i}^{N}\exp\Big(\mathrm{sim}\Big(Z_i^{(1)}, Z_k^{(2)}\Big)/T\Big)}, \quad (12)$$

where T denotes the temperature parameter, and $\mathrm{sim}(\cdot)$ represents the similarity function.

The overall optimization objective of our model can be defined as:

$$\mathcal{L} = \mathcal{L}_c + \alpha\mathcal{L}_{cd} + \beta\mathcal{L}_{co}, \quad (13)$$

where α is a balancing parameter that controls the trade-off between different loss terms, and β regulates the influence of pseudo-label refinement during the overall training process.

III. Experiment

TABLE I
THE STATISTICS AND PARAMETER SETTINGS OF THE DATASETS

Dataset	Dataset Statistics				Params	
	Samples	Dimension	Edges	Classes	α	β
CORA	2708	1433	5429	7	0.5	0.4
AMAP	7650	745	119081	8	0.5	0.3
UAT	1190	239	13599	4	0.5	0.4
BAT	131	81	1038	4	0.5	0.4
EAT	399	203	5994	4	0.8	0.3

A. Experimental Setting

The experiments are conducted on a machine equipped with an Intel i9-14900KF CPU, an NVIDIA RTX 4090 GPU, and 256 GB of RAM. For all datasets, the model is trained for 400 epochs using the Adam optimizer. We utilized five widely-used datasets in our experiments: CORA, AMAP, BAT, EAT and UAT. The learning rates for training are set as follows: 5e-4 for Cora, 5e-2 for AMAP, 1e-4 for BAT, 5e-3 for UAT, and 1e-5 for EAT. Clustering performance is assessed using four widely adopted metrics: ACC, NMI, ARI, and F1. The statistics and hyper-parameter settings are summarized in Table I .

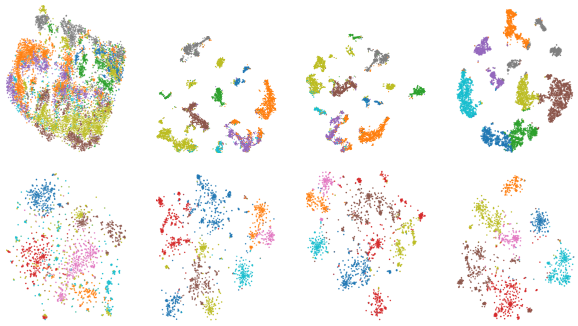
B. Experimental Result

We conduct comparative experiments on five public graph clustering benchmark datasets, where the proposed method is evaluated against eleven baseline approaches, as shown in Table II. Specifically, the compared methods can be broadly categorized into three groups: classical graph clustering methods (SDCN [2]), graph contrastive clustering methods (MVGRL [6],AGC-DRR [3],DCRN [14],HSAN [8],HomoCAGC [11]), and graph augmentation-based clustering methods (AutoSSL [12],NCLA [7],CONVERT [16],GraphLearner [17],SECL [10]).

On the AMAP dataset, our method achieves the best performance across all four metrics. In particular, it attains ACC, NMI, and ARI scores of 80.84%, 72.04%, and 63.41%, respectively, outperforming the strongest baseline HomoCAGC

TABLE II
CLUSTERING RESULTS ON FIVE DATASETS. BEST RESULTS ARE BOLD VALUES AND THE SECOND BEST VALUES ARE UNERLINED.

Dataset	Metric	SDCN	MVGRL	AGC-DRR	AutoSSL	DCRN	NCLA	HSAN	CONVERT	SECL	GraphLearner	HomoCAGC	Ours
CORA	ACC	35.60	70.47	40.62	63.81	69.93	51.09	77.07	74.07	74.79	74.91	76.77	**77.62**
	NMI	14.28	55.57	18.74	47.62	45.13	31.80	59.21	55.57	56.88	58.16	**60.17**	59.06
	ARI	07.78	48.70	14.80	38.92	33.15	36.66	57.52	50.58	52.84	53.82	**58.37**	56.30
	F1	24.37	67.15	31.23	56.42	49.50	51.12	75.11	72.92	73.29	73.33	74.46	**76.61**
AMAP	ACC	53.44	41.07	76.81	54.55	76.39	67.18	77.02	77.19	77.76	77.24	77.44	**80.84**
	NMI	44.85	30.28	66.54	48.56	67.38	63.63	67.21	67.20	67.67	67.12	69.21	**72.04**
	ARI	31.21	18.77	60.15	26.87	57.56	46.30	58.01	60.79	58.64	58.14	58.43	**63.41**
	F1	50.66	32.88	71.03	54.47	70.98	73.04	72.03	74.03	72.29	73.02	71.72	**79.88**
BAT	ACC	53.05	37.56	47.79	42.43	67.94	47.48	77.15	78.02	78.40	75.50	77.86	**80.61**
	NMI	25.74	29.33	19.91	17.84	47.23	24.36	53.21	53.54	53.74	50.58	54.43	**57.09**
	ARI	21.04	13.45	14.59	13.11	39.76	24.14	52.20	51.95	51.52	47.45	52.61	**56.14**
	F1	46.45	29.64	42.33	34.84	67.40	42.25	77.13	77.77	78.47	75.40	77.60	**80.49**
EAT	ACC	39.07	32.88	37.37	31.33	50.88	36.06	56.69	58.35	58.00	57.22	53.38	**58.40**
	NMI	08.83	11.72	07.00	07.63	22.01	21.46	33.25	33.36	33.98	33.47	**35.78**	33.66
	ARI	06.31	04.68	04.88	02.13	18.13	21.48	26.85	27.11	27.25	26.21	24.48	**27.34**
	F1	33.42	25.35	35.20	21.82	47.06	31.25	57.26	58.42	58.15	57.53	54.23	**58.51**
UAT	ACC	52.25	44.16	42.64	42.52	49.92	45.38	56.04	57.36	58.28	55.31	56.72	**58.71**
	NMI	21.61	21.53	11.15	17.86	24.09	24.49	26.99	28.75	**28.96**	24.40	28.39	27.51
	ARI	21.63	17.12	09.50	13.13	17.17	21.34	25.22	27.96	25.39	22.14	21.50	**28.11**
	F1	45.59	39.44	35.18	52.94	44.81	30.56	54.20	54.55	56.66	52.77	56.96	**57.03**



(a) epoch=0 (b) epoch=150 (c) epoch=300 (d) epoch=400

Fig. 2. 2D visualization of embeddings at different training epochs on the AMAP (top) and CORA (bottom) datasets.

by 3.40%, 2.83%, and 4.98%. The improvement in ARI indicates that the learned representations better capture cluster boundaries in the AMAP dataset. Similar performance gains are observed on the BAT datasets. On BAT, our method achieves ACC and F1 scores of 80.61% and 80.49%, surpassing HomoCAGC by 2.75% and 2.89%, respectively. On EAT, despite the overall difficulty of the dataset, our method attains the best ARI score of 27.34%. Overall, the proposed model demonstrates consistent and competitive performance across multiple benchmark datasets and evaluation metrics, validating its effectiveness for graph contrastive clustering.

C. Visualization Analysis

We visualize the learned graph embeddings Z using t-SNE to qualitatively assess the clustering performance (Fig. 2). We conduct the analysis on the AMAP and CORA datasets. As training progresses, nodes from different classes gradually evolve from an initially intertwined distribution to a more clearly separated layout, while nodes within the same class become increasingly compact. This observation indicates that the proposed model progressively captures more fine-grained class information.

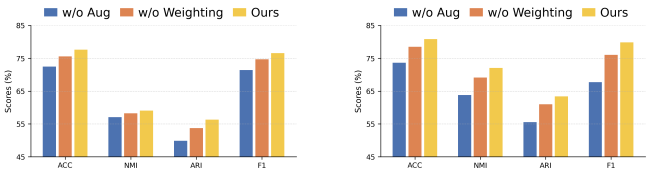
TABLE III
ABLATION EXPERIMENTS ON LOSSES ACROSS FIVE BENCHMARK DATASETS. THE BEST RESULTS HAVE BEEN MARKED IN BOLD.

Dataset	**Metric**	w/o $\mathcal{L}_{cd}$	w/o $\mathcal{L}_{c}$	w/o $\mathcal{L}_{co}$	**UGCC**
CORA	ACC	76.89	71.21	76.89	**77.62**
	NMI	58.25	52.58	58.52	**59.06**
	ARI	54.75	46.25	55.43	**56.30**
	F1	76.22	71.06	75.90	**76.61**
AMAP	ACC	79.28	74.78	78.34	**80.84**
	NMI	69.73	62.20	68.65	**72.04**
	ARI	61.50	54.39	60.26	**63.41**
	F1	77.09	71.47	75.52	**79.88**
BAT	ACC	79.16	78.24	80.31	**80.61**
	NMI	55.15	54.93	56.80	**57.09**
	ARI	53.58	52.85	55.69	**56.14**
	F1	79.10	77.90	80.13	**80.49**
EAT	ACC	58.25	58.05	58.02	**58.35**
	NMI	33.35	33.23	33.48	**33.66**
	ARI	27.13	26.99	27.07	**27.34**
	F1	58.32	58.09	58.14	**58.51**
UAT	ACC	58.65	55.33	58.46	**58.71**
	NMI	26.47	27.13	27.43	**27.51**
	ARI	24.93	23.53	27.21	**28.11**
	F1	56.29	53.12	56.68	**57.03**

D. Ablation Studies

In this subsection, we conduct ablation studies to evaluate the contribution of each loss component. Specifically, we denote $\mathcal{L}_{cd}$, $\mathcal{L}_{c}$, and $\mathcal{L}_{co}$ as the variants without consistency–diversity balancing loss, contrastive loss, and uncertainty-aware clustering alignment loss, respectively. Table III summarizes the results. Overall, removing any loss term degrades performance to varying degrees. Among all variants, removing $\mathcal{L}_{c}$ causes the largest drop. On CORA, ACC decreases from 77.62% to 71.21% and ARI drops from 56.30% to 46.25%. This highlights the critical role of the contrastive loss in learning discriminative

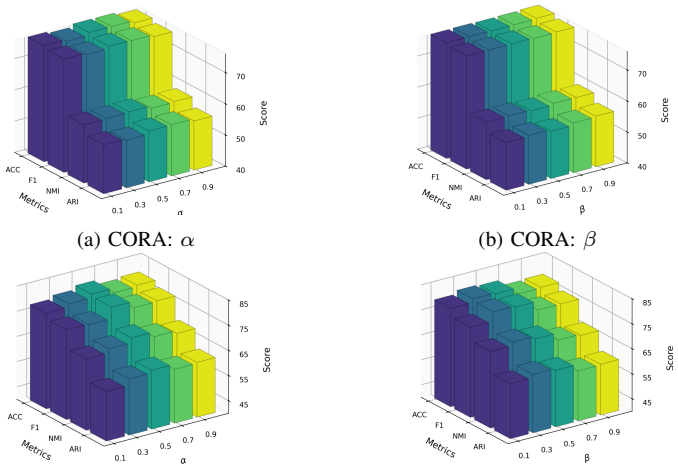
representations and preserving clustering structure. Removing $\mathcal{L}_{cd}$ also degrades performance. For example, on AMAP, ACC and NMI decrease from 80.84% to 79.28% and from 72.04% to 69.73%, respectively, indicating that $\mathcal{L}_{cd}$ helps stabilize the learnable augmentation process. Excluding $\mathcal{L}_{co}$ results in consistent performance drops, particularly on ARI. For instance, ARI decreases from 56.14% to 55.69% on BAT and from 56.30% to 55.43% on CORA, demonstrating its effectiveness in aligning clustering structures across views.



(a) CORA (b) AMAP

Fig. 3. Ablation results of the proposed module on CORA and AMAP.

We further evaluate the proposed coupled dual-path augmentation module and uncertainty-aware weighting by removing each component. Here, (w/o) Aug denotes the variant without the dual-path augmentation module, and (w/o) Weighting denotes the variant without uncertainty-aware weighting. As shown in Fig. 3, incorporating all components consistently improves performance on CORA and AMAP, validating the effectiveness of our design.



(a) CORA: α (b) CORA: β

(c) AMAP: α (d) AMAP: β

Fig. 4. Impact of parameters α and β on the CORA and AMAP datasets.

E. Hyper-parameter Analysis

we analyze the parameters α and β, which control the augmentation regularization term and the uncertainty-aware clustering alignment term, respectively. With other settings fixed, we perform a coarse-grained search over [0.1,0.9] on CORA and AMAP. As shown in Fig 4, UGCC is not sensitive to these hyperparameters and achieves stable results.

IV. Conclusion

In this paper, we propose UGCC, a graph contrastive clustering method. We design a coupled dual-path augmentation module that introduces learnable and adaptive augmentation from both attribute and structural perspectives, generating multi-view samples that are better aligned with clustering semantics. To further couple representation learning with clustering objectives, we develop an uncertainty-aware clustering alignment strategy that adaptively weights clustering constraints based on prediction uncertainty, stabilizing unsupervised optimization. In addition, a mutually balanced objective is introduced to guarantee the consistency of the embeddings as well as the diversity of the augmented samples. Extensive experiments on five datasets demonstrate the effectiveness and competitiveness of UGCC.

References

[1] Z. Zhou, B. Tang, F. Chen, W. Wang, S. Zhao, and N. Yu, "scscdt: Self-contrastive neural network with deep topology mining for scrna-seq data clustering," *Expert Systems with Applications*, p. 129751, 2025.

[2] D. Bo, X. Wang, C. Shi, M. Zhu, E. Lu, and P. Cui, "Structural deep clustering network," in *Proceedings of the web conference 2020*, 2020, pp. 1400–1410.

[3] L. Gong, S. Zhou, W. Tu, and X. Liu, "Attributed graph clustering with dual redundancy reduction." in *IJCAI*, 2022, pp. 3015–3021.

[4] J. Xie, R. Girshick, and A. Farhadi, "Unsupervised deep embedding for clustering analysis," in *International conference on machine learning*. PMLR, 2016, pp. 478–487.

[5] J. Yin, H. Wu, and S. Sun, "Effective sample pairs based contrastive learning for clustering," *Information Fusion*, vol. 99, p. 101899, 2023.

[6] K. Hassani and A. H. Khasahmadi, "Contrastive multi-view representation learning on graphs," in *International conference on machine learning*. PMLR, 2020, pp. 4116–4126.

[7] X. Shen, D. Sun, S. Pan, X. Zhou, and L. T. Yang, "Neighbor contrastive learning on learnable graph augmentation," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, no. 8, 2023, pp. 9782–9791.

[8] Y. Liu, X. Yang, S. Zhou, X. Liu, Z. Wang, K. Liang, W. Tu, L. Li, J. Duan, and C. Chen, "Hard sample aware network for contrastive deep graph clustering," in *Proc. of AAAI*, 2023.

[9] Y. Zhao and L. Bai, "Contrastive clustering with a graph consistency constraint," *Pattern Recognition*, vol. 146, p. 110032, 2024.

[10] X. Wu, J. Hu, A. Zhang, Y. Quan, Q. Miao, and P. G. Sun, "Structure-enhanced contrastive learning for graph clustering," *IEEE Transactions on Network Science and Engineering*, 2025.

[11] M.-S. Chen, P.-Y. Lai, D.-Z. Liao, C.-D. Wang, and J.-H. Lai, "Homophily induced contrastive attributed graph clustering," *IEEE Transactions on Circuits and Systems for Video Technology*, 2025.

[12] W. Jin, X. Liu, X. Zhao, Y. Ma, N. Shah, and J. Tang, "Automated self-supervised learning for graphs," *arXiv preprint arXiv:2106.05470*, 2021.

[13] H. Zhao, X. Yang, Z. Wang, E. Yang, and C. Deng, "Graph debiased contrastive learning with joint representation clustering." in *IJCAI*, 2021, pp. 3434–3440.

[14] Y. Liu, W. Tu, S. Zhou, X. Liu, L. Song, X. Yang, and E. Zhu, "Deep graph clustering via dual correlation reduction," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 36, no. 7, 2022, pp. 7603–7611.

[15] S. Zhao, B. Tang, Z. Zhou, N. Yu, and F. Chen, "Dual contrastive attributed graph clustering with adaptive structure," in *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*. Springer, 2025, pp. 470–483.

[16] X. Yang, C. Tan, Y. Liu, K. Liang, S. Wang, S. Zhou, J. Xia, S. Z. Li, X. Liu, and E. Zhu, "Convert: Contrastive graph clustering with reliable augmentation," in *Proceedings of the 31st ACM international conference on multimedia*, 2023, pp. 319–327.

[17] X. Yang, E. Min, K. Liang, Y. Liu, S. Wang, S. Zhou, H. Wu, X. Liu, and E. Zhu, "Graphlearner: Graph node clustering with fully learnable augmentation," in *Proceedings of the 32nd ACM international conference on multimedia*, 2024, pp. 5517–5526.