CLASS CENTROID-GUIDED ALIGNMENT FOR MULTI-VIEW CROSS DOMAIN GRAPHS

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-view graph domain adaptation (MGDA) addresses label scarcity by transferring knowledge from a labeled source graph to an unlabeled target graph, where nodes are connected through multiple views. Most existing methods align each view as a whole and ignore class-level structure, so that classes can remain misaligned across domains even when view-level distributions match. We empirically reveal that the cross-domain discrepancy in class-level structure is strongly associated with alignment quality. To capture this structure, we use each view's class centroids as anchors and leverage the Cross-Domain Aggregated Class Distance (CD-ACD), which compares source and target inter-class centroid distances within each view and quantifies each view's class-level structural shift. Based on CD-ACD, we propose Class Centroid-Guided Alignment (CCGA), which uses the metric to weight the alignment of each view and class pair during training and thereby reduces class-level structural shifts between domains in each view. Experimental results on various multi-view graph benchmark datasets demonstrate the effectiveness of our method.

1 INTRODUCTION

Multi-view graph contains node attributes and multiple relations, to represent complex systems, so they are widely applied in scenarios such as citation networks, social networks, and knowledge graphs (Qu et al., 2017; Wu et al., 2023b; Pan & Kang, 2023). However, manually annotating such graph data with complex topological structures is extremely difficult and highly expensive, leading to a severe issue of label scarcity (Huang et al., 2024; Zhuo et al., 2026). To address this problem, Multi-view Graph Domain Adaptation (MGDA) has emerged as an effective paradigm, aiming to transfer knowledge from a label-rich source graph to an unlabeled target graph, where nodes are connected through multiple views (Fang et al., 2026).

Despite the immense potential of MGDA, cross-domain transfer still faces significant challenges, since single-view graphs and multi-view graphs possess more complex structures, sepecially the structural differences formed by the nodes from different classes. Most existing MGDA methods align each view as a whole. For instance, SAGA (Fang et al., 2026) extracts topological information from both view and hop perspectives for graph alignment, yet it fails to account for class-level structure. However, if alignment is only achieved at the view-level distribution, the same-class nodes across domains may remain misaligned, thereby degrading the quality of cross-domain alignment.

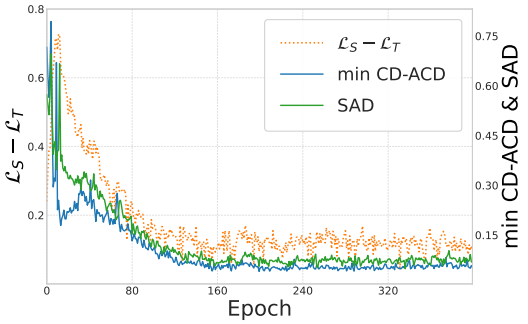


Figure 1: This represents the training dynamics curves of CD-ACD on the ACM dataset. This shows a strong correlation between class-level structural information and alignment quality.

To verify our argument, we empirically reveal that cross-domain discrepancy in class-level structure is strongly associated with alignment quality. To capture this structural feature, we design Cross-Domain Aggregated Class Distance (CD-ACD) by

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.

utilizing each view's class centroids as anchors, and the centroid distances of the source and target classes within each view to quantify each view's class-level structural shift. As illustrated in Figure 1, it can be observed that as the training epochs increase, the cross-domain classification loss discrepancy ($\mathcal{L}_S - \mathcal{L}_T$) represented by the left y-axis progressively decreases, the value of min CD-ACD from our method and Structural Aggregation Distance (SAD) metric from SAGA method (Fang et al., 2026) decreases correspondingly, and min CD-ACD converges to a lower discrepancy value than SAD. Based on this observation, we can conclude that a smaller CD-ACD indicates that source domain and target domain achieve higher-quality cross-domain alignment at class level, and CD-ACD metric is more effective than the existing SAD metric. More detailed empirical study results on all datasets are provided in Appendix F.

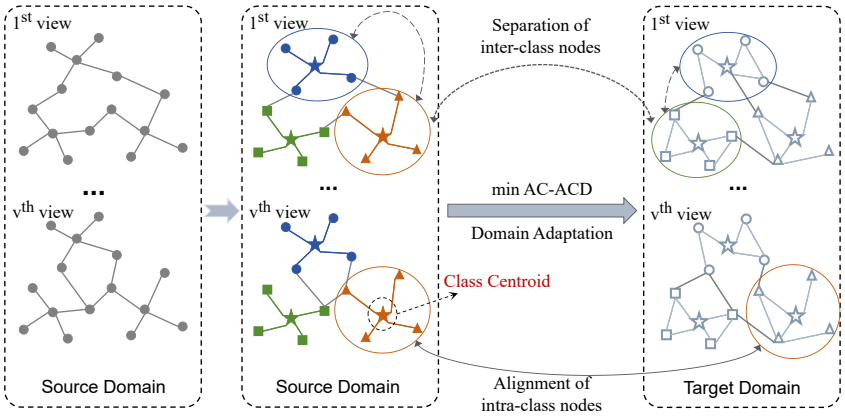


Figure 2: Illustration of class centroids and cross-domain alignment.

Motivated by this observation, we propose Class Centroid-Guided Alignment framework (CCGA), which dynamically guides cross-domain alignment for multi-view graphs between source domain and target domain. Specifically, CCGA utilizes CD-ACD metric during training to dynamically weight the alignment process for each view and class pair. CCGA captures class-level structural differences across views and dynamically adjusts its alignment strategy, and aligns view-level topological information, so as to reduce the class-level structural shifts between source domain and target domain within each view, achieving higher-quality cross-domain alignment.

The main contributions of this paper are summarized as follows:

- We empirically reveal that cross-domain discrepancy in class-level structure is strongly associated with alignment quality, and design the metric Cross-Domain Aggregated Class Distance (CD-ACD) to quantify each view's class-level structural shift.
- We propose Class Centroid-Guided Alignment method (CCGA) for multi-view graph adaptation, which utilizes CD-ACD metric to dynamically weight and guide the alignment process of view and class pairs, and mitigates the issue of cross-domain class-level structural shift across domains.
- Comprehensive experiments on benchmarks show the superior performance of our method compared to other state-of-the-art methods for real-world datasets of the cross-network node classification tasks.

2 RELATED WORK

2.1 MULTI-VIEW GRAPH LEARNING

Unsupervised multi-view learning aims to address unsupervised tasks in multiplex graphs by leveraging node features and graph structures. Early multi-view graph learning methods MvAGC (Lin et al., 2021) and MCGC (Pan & Kang, 2021) combine graph filtering with unsupervised techniques such as spectral clustering and subspace clustering to reveal underlying patterns in complex networks. With the rise of deep representation learning, the unsupervised learning of low-dimensional node

representations methods strat to utilize Graph Neural Networks and self-supervised techniques has emerged as a new paradigm. O2MAC (Fan et al., 2020) uses GNNs in UMGL to capture shared information by selecting the most informative graph and reconstructing all graph structures. DMGI (Park et al., 2020) and HDMI (Jing et al., 2021) fuse representations from different relations by maximizing the mutual information between local and global contexts. MGDCR (Mo et al., 2023a) and BTGF (Qian et al., 2024) employ various contrastive losses to align representations from different relations, thereby preventing dimensional collapse. CoCoMG (Peng et al., 2023) and DMG (Mo et al., 2023b) capture comprehensive information by learning inter-graph consistency and complementarity. InfoMGF (Shen et al., 2024b) propose an information-aware unsupervised multiplex graph fusion framework to maximizes view-shared and view-specific task-relevant information while eliminating irrelevant noise. To address the partitioning bottleneck of large-scale multi-view data, RCAL (Qin et al.) proposes a robust consensus anchor learning method for multi-view subspace clustering, which significantly reduces computational complexity and enhances model scalability. CoE (Wang et al., 2025b) proposes a collaborative expert mechanism, which integrates Graph Mixture of Experts (Graph-MoE) model with large margin optimization, to address the challenges of heterogeneous information fusion and knowledge transfer in multimodal and multi-view graph data.

2.2 Graph Domain Adaptation

Unsupervised Graph Domain Adaptation (GDA) aims to transfer knowledge from a label-rich source domain graph to an unlabeled target domain graph. Recent studies have explored advanced generative and discriminative paradigms for feature distribution alignment (Zhang et al., 2026; Luo et al., 2025; Yin et al., 2025). ADA-NCF (Chen et al., 2026a) and GMM-GDA (Wang et al., 2025a) utilize neural characteristic functions and Gaussian mixture models to achieve more precise distribution matching on GDA graphs. Furthermore, several studies extend graph alignment to multi-source scenarios through prototypical contrastive learning and hierarchical structures (Shen et al., 2026; Lin et al., 2025). RCS (Luo et al., 2026) rely on pre-trained models and unlabeled target graphs to address data privacy constraints (Luo et al., 2026; 2024; Zhang et al., 2024). DPEC (Ma et al., 2026) tackles the class distribution mismatch problem in GDA via dual prototype enhancement, dual reprogramming, and mixup exploration (Ma et al., 2026; 2025; Zhang & He, 2026; Yin et al., 2024). IFI (Qiao et al., 2024; 2023) leverages information filtering and interpolating to maximize the utility of limited target labels.

A core challenge in GDA is to identify and mitigat specific structural discrepancies, such as conditional structure shifts and homophily shifts. To overcome topology and label shifts, recent studies adaptively recalibrate edge weights alongside graph structural re-weighting. SSET (Chen et al., 2026b) and DGSDA (Yang et al., 2025) explore structure-semantic evolution trajectories and disentangled spectral domain adaptation. PGSA (Wen et al.) and HGDA (Fang et al., 2025) mitigate domain discrepancies by explicitly aligning homophily, heterophily, and attribute signals via dedicated filters. Furthermore, SAGA framework (Fang et al., 2026) introduces Structural Aggregation Distance to guide the alignment of multi-view graphs by dynamically selecting neighborhood orders, so as to effectively characterize and mitigate view-level and hop-level structural shifts.

3 Cross-Domain Aggregated Class Distance

Notations. Source domain multi-view graph $G^S = \{\mathcal{V}^{\mathcal{S}}, \{\mathcal{E}_v^S\}_{v=1}^V, \{A_v^S\}_{v=1}^V, X^S, Y^S\}$ and target domain multi-view graph $G^T = \{\mathcal{V}^{\mathcal{T}}, \{\mathcal{E}_v^T\}_{v=1}^V, \{A_v^T\}_{v=1}^V, X^T\}$. The domain $D \in \{S, T\}$ graph, consist of a set of nodes $\mathcal{V}^D$ and view-specific edges $\{\mathcal{E}_v^D\}_{v=1}^V$, along with a set of adjacency matrices $\{A_v^D\}_{v=1}^V$, a feature matrix X^D, and a label matrix Y^S. For each view v, the adjacency matrix $A_v^D \in \mathbb{R}^{N^D \times N^D}$ encodes the connections between N nodes. The feature matrix $X^D \in \mathbb{R}^{N^D \times d}$ represents node features, with each node described by a d-dimensional feature vector. $Y^S \in \mathbb{R}^{N^S \times C}$ contains the labels for N nodes, where each node is classified into one of C classes. The numbers of source domain nodes and target domain notes are denoted as N^S and N^T respectively, and the total number of views in multi-relational graph is denoted as V.

Existing approaches struggle to measure fine-grained class-level structural shifts within unlabeled target domains. To address this limitation, we propose Cross-Domain Aggregated Class Distance (CD-ACD), to characterize structural deviations through explicit class centroids and normalized

distance components, CD-ACD quantifies multi-view structural alignment discrepancies under domain shifts and dynamically evaluates the alignment quality of individual views.

In this formulation, we calculate class centroids across varying views. For the source domain with the label matrix Y^S, we apply a hard aggregation mechanism; For the unlabeled target domain, we extract the prediction probability matrix $P^T \in [0,1]^{N^T \times C}$, where the element $P^T_{j,m}$ denotes the predicted probability of target node j belonging to class m. With these predicted probabilities, we formulate a dynamic confidence weight $W^T_{j,m}$ to signify the confidence of target node j belonging to class m. For a given target node j, the weight $W^T_{j,m}$ is determined by incorporating a truncation threshold τ_{conf} and an exponential smoothing parameter γ_{pow} to filter out low-confidence label assignments, defined as follows:

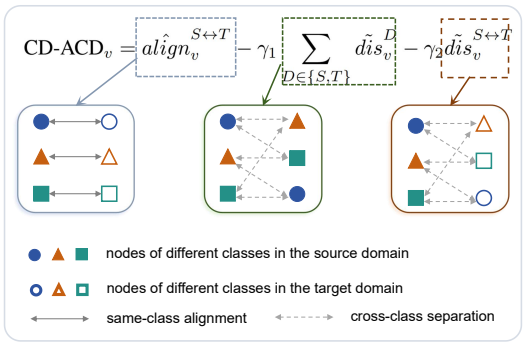


Figure 3: The components of Cross-Domain Aggregated Class Distance.

$$W^T_{j,m} = P^T_{j,m} \cdot (\max_{c\in\{1,\dots,C\}} P^T_{j,c})^{\gamma_{pow}} \cdot \mathbb{I}(\max_{c\in\{1,\dots,C\}} P^T_{j,c} \geq \tau_{conf}). \tag{1}$$

The class centroids $\bar{z}$ for the m-th class in the source and target domains under view v are respectively defined as:

$$\bar{z}^S_{m,v} = \frac{\sum_{i=1}^{N^S} Y^S_{i,m} z^S_{i,v}}{\sum_{i=1}^{N^S} Y^S_{i,m}}, \quad \bar{z}^T_{m,v} = \frac{\sum_{j=1}^{N^T} W^T_{j,m} z^T_{j,v}}{\sum_{j=1}^{N^T} W^T_{j,m}}, \tag{2}$$

where $z^T_{j,v}$ denotes the feature representation of node j under view v in target domain.

CD-ACD is constructed by three normalized distance components: cross-domain same-class alignment distance $\hat{align}_v^{S\leftrightarrow T} = \frac{1}{C}\sum_{m=1}^{C} ||\bar{z}^S_{m,v} - \bar{z}^T_{m,v}||_2$, intra-domain class separation distance $\tilde{dis}_v^D = \frac{2}{C(C-1)}\sum_{m=1}^{C}\sum_{n=m+1}^{C} ||\bar{z}^D_{m,v} - \bar{z}^D_{n,v}||_2$, and cross-domain cross-class separation distance $\tilde{dis}_v^{S\leftrightarrow T} = \frac{1}{C(C-1)}\sum_{m=1}^{C}\sum_{n\neq m}^{C} ||\bar{z}^S_{m,v} - \bar{z}^T_{n,v}||_2$. Furthermore, we utilize a multi-view mean normalization operator to mitigate scale disparities among different distance evaluation terms.

The final CD-ACD score for view v is derived from a linear combination of these components, governed by balancing hyperparameters γ_1 and γ_2:

$$\text{CD-ACD}_v = \hat{align}_v^{S\leftrightarrow T} - \gamma_1 \sum_{D\in\{S,T\}} \tilde{dis}_v^D - \gamma_2 \tilde{dis}_v^{S\leftrightarrow T}. \tag{3}$$

4 METHODOLOGY

In this section, we present CCGA method for multi-view cross domain graphs. As illustrated in Figure 4, our methodology follows a progressive design: (i) decouple graph propagation and feature mapping to obtain view-specific topological features and low-dimensional representations, (ii) construct class centroids to calculate CD-ACD, so as to guide joint cross-domain feature alignment dynamically, and (iii) concatenate multi-view features to construct global node representations for target node classification and achieve cross-domain target node classification.

4.1 SCALABLE GRAPH FEATURE RECONSTRUCTION

To enhance model scalability on multi-view graphs, we adopt Approximate Personalized Propagation of Neural Predictions (APPNP) architecture to strictly decouple graph propagation and feature mapping processes. The model independently performs iterative aggregation via Personalized PageRank on each view to obtain view-specific topological features. For domain $D \in \{S,T\}$, starting from the original features X^D, the propagation process under view v is formulated as:

$$X^D_{v,0} = X^D, \quad X^D_{v,k+1} = (1-\alpha)\hat{A}^D_v X^D_{v,k} + \alpha X^D, \tag{4}$$

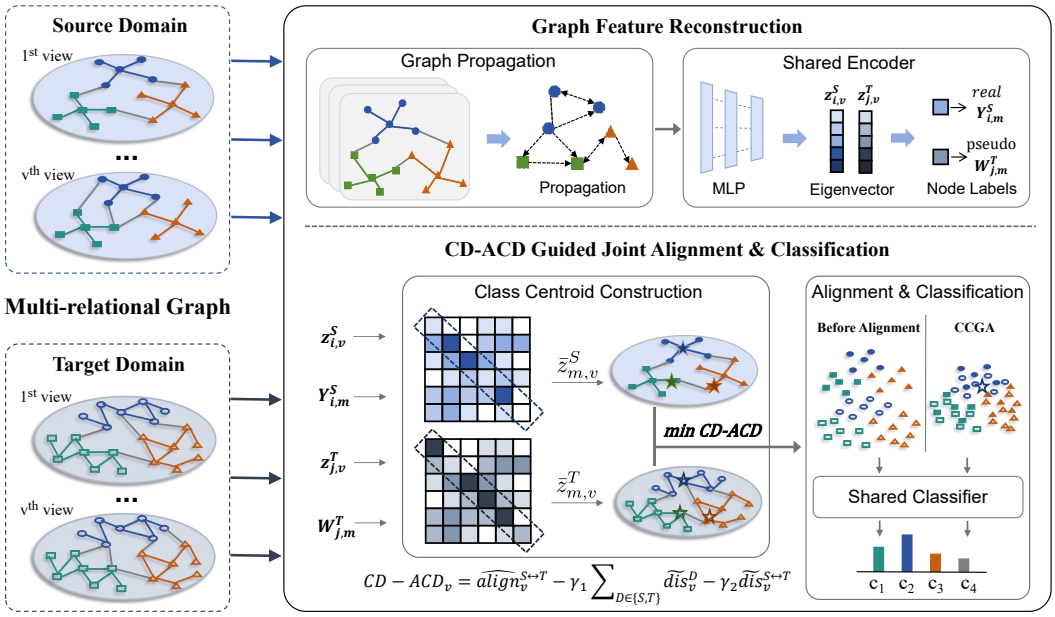


Figure 4: Overview of CCGA framework. It performs graph feature reconstruction to obtain node representations, calculates the CD-ACD based on constructed class centroids, and dynamically guides joint cross-domain alignment and target node classification.

where $\hat{A}_v^D$ is the normalized adjacency matrix with added self-loops, and the hyper-parameter $\alpha \in [0, 1]$ represents the teleport probability..

After K steps of propagation, the smoothed aggregated features $X_{v,(K)}^D$ is obtained, and then directly feed the into a shared non-linear feature encoder $f_\Theta(\cdot)$, and mapped into low-dimensional latent representations Z_v^D:

$$Z_v^D = f_\Theta(X_{v,(K)}^D). \tag{5}$$

Subsequently, the model utilizes a shared decoder $g_\Theta(\cdot)$ to reconstruct the aggregated features, ensuring that the dimension-reduced representations preserve the essential structure and attribute information of the graph. The multi-view feature reconstruction loss $\mathcal{L}_R$ applies a global constraint by calculating Mean Squared Error (MSE), quantified by the squared Frobenius norm $||\cdot||_F^2$, between the decoded features $g_\Theta(Z_v^D)$ and the original aggregated features across all V views for both source domain and target domain:

$$\mathcal{L}_R = \frac{1}{V}\sum_{v=1}^{V}\left(\frac{1}{N^S \cdot d}||g_\Theta(Z_v^S) - X_{v,(K)}^S||_F^2 + \frac{1}{N^T \cdot d}||g_\Theta(Z_v^T) - X_{v,(K)}^T||_F^2\right), \tag{6}$$

where the encoder $f_\Theta(\cdot)$ and decoder $g_\Theta(\cdot)$ are both implemented by Multilayer Perceptron (MLP).

4.2 Aggregation Class Distance-Guided Joint Alignment

The model leverages Cross-Domain Aggregated Class Distance to dynamically explore unsupervised view alignment strategies. To overcome the limitation of isolating cross-domain feature alignment from intra-domain consistency, we propose an alignment mechanism guided by CD-ACD metric.

We calculate CD-ACD by quantifying the structural shift between the source and target domain under each view. Then Softmax confidence weight ω_v is designed accoding to CD-ACD score with a temperature parameter λ to adaptively evaluate the multi-view alignment status, so as to guide feature aggregation across multiple views:

$$\omega_v = \frac{\exp(-\lambda \cdot \text{CD-ACD}_v)}{\sum_{j=1}^{V}\exp(-\lambda \cdot \text{CD-ACD}_j)}. \tag{7}$$

In the dynamic weight distribution, the minimum CD-ACD score corresponds to the maximum weight parameter, indicating that the class topological structure shift is smaller under this specific view between source domain and target domain. We denote the weight corresponding to the minimum CD-ACD value as $\omega_b = \max(\omega_v)$, and utilize CD-ACD$_b$ jointly with the weight set ω_v to guide the cross-domain alignment process.

To achieve cross-domain space alignment, a non-linear projection $\psi(\cdot)$ is integrated into the view feature space corresponding to the minimum CD-ACD where bidirectional InfoNCE contrastive loss is optimized to the across domains representation distance. The unidirectional cross-domain contrastive loss is defined as follows:

$$l(Z_k^S, Z_k^T) = -\frac{1}{N^S}\sum_{i=1}^{N^S}\log\left[\frac{1}{N^S N^T}\sum_{j=1}^{N^T}\exp\left(\frac{\psi(Z_{i,k}^S)\cdot\psi(Z_{j,k}^T)^\top}{\tau\cdot||\psi(Z_{i,k}^S)||_2||\psi(Z_{j,k}^T)||_2}\right)\right]. \tag{8}$$

Taking the first-order graph propagation view ($v = 1$) as the baseline anchor, the model utilizes ω_v to dynamically penalize the Frobenius norm differences of the covariance matrices of each view in source domain and target domain ($D \in \{S, T\}$). By introducing the weight coefficient $\omega_b = \max(\omega_v)$, the model constructs the joint alignment loss $\mathcal{L}_A$ based on CD-ACD weights:

$$\mathcal{L}_A = \omega_b\Big[l(Z_k^S, Z_k^T)+l(Z_k^T, Z_k^S)\Big]+\frac{1-\omega_b}{V}\sum_{v=1}^{V}\sum_{D\in\{S,T\}}\omega_v\left|\left|\frac{1}{N^D}(Z_v^D)^\top Z_v^D - \frac{1}{N^D}(Z_1^D)^\top Z_1^D\right|\right|_F^2. \tag{9}$$

We get CD-ACD to quantify the structural shift of each view between source domain and target domain, and guide high-quality cross-domain feature alignment dynamically.

4.3 Target Node Classification

The low-dimensional features of each sub-view are obtained to concatenate along the feature dimension to construct unified node representation $Z_U^D = \text{Concat}(Z_1^D, ..., Z_V^D)$ incorporated the global multi-view information. We utilize a linear classifier $h_\Phi(\cdot)$ to map this global representation into the class space, and calculate the predicted log-probabilities for each class as $O^D = h_\Phi(Z_U^D)$.

For source domain data with ground truth labels, the standard cross-entropy loss $\mathcal{L}^S$ is employed to optimize the classifier, extracting class-discriminative features from source domain:

$$\mathcal{L}_S = -\frac{1}{N^S}\sum_{i=1}^{N^S}\sum_{c=1}^{C}Y_{i,c}^S\log\left(\frac{\exp(O_{i,c}^S)}{\sum_{k=1}^{C}\exp(O_{i,k}^S)}\right), \tag{10}$$

where $O_{i,c}^D$ denotes the predicted log-probability output by the linear classifier for node i in domain D ($D \in \{S, T\}$) belonging to class c, and C is the total number of classes.

For the unlabeled target domain data, minimizing the target domain classification information entropy $\mathcal{L}^T$ pushes the classification decisions away from the dense data classes of the target domain, so as to enhance the model's unsupervised confidence in target node classification:

$$\mathcal{L}_T = -\frac{1}{N^T}\sum_{j=1}^{N^T}\sum_{c=1}^{C}P_{j,c}^T\log\left(\frac{\exp(O_{j,c}^T)}{\sum_{k=1}^{C}\exp(O_{j,k}^T)}\right). \tag{11}$$

The overall optimization objective function is formulated combining all the aforementioned modules, including the multi-view feature reconstruction loss $\mathcal{L}_R$, the joint alignment loss $\mathcal{L}_A$, the source domain classification loss $\mathcal{L}^S$, and the target domain information entropy optimization loss $\mathcal{L}^T$. During the training process, the model jointly optimizes the overall loss through backpropagation:

$$\mathcal{L} = \mathcal{L}_R + \mathcal{L}_A + \beta\mathcal{L}^S + \delta\mathcal{L}^T, \tag{12}$$

where β and δ are hyperparameters used to balance the contribution weights of each term. A detailed description of our proposed CCGA method is provided in Appendix I.

5 EXPERIMENTS

5.1 DATASETS

We conduct experiments on real-world citation network datasets to demonstrate the effectiveness on domain adaptation node classification. In real-world citation network graphs, nodes correspond to papers, and edges denote the citation relationships between them. The edges are defined by two types of meta-paths: ACM1 and ACM2 include PAP (Paper-Author-Paper) and PSP (Paper-Subject-Paper), whereas MAG contains PAP (Paper-Author-Paper) and PP (Paper-Paper).

ACM1 and ACM2 graphs are labeled into three classes: Database, Wireless Communication, and Data Mining. MAG is the subset of OGBN-MAG (Wang et al., 2020), which is partitioned into six distinct national subgraphs (CN, US, DE, FR, JP, and RU), with the ten most frequent classes extracted. In multi-view graph node classification experiments, any two national subgraphs can serve as source dataset and target dataset respectively. More details regarding the dataset statistics, baseline methods, and the specific experimental setup can be found in Appendix A, Appendix B, and Appendix C.

5.2 EXPERIMENTAL SETUP

Our multi-view graph setting contains multiple views for both source domain and target domain. To ensure a fair comparison, all baselines are evaluated under a consistent training/evaluation protocol, with single-view baselines explicitly addressed within multi-view environment.

For non-adaptive baselines (e.g., GCN (Kipf & Welling, 2016), GAT (Veličković et al., 2017)), each view in source domain is independently processed according to its original formulation to obtain view-specific embeddings. We then average the embeddings across all source views and utilize these averaged embeddings along with the available labels to train GNN. We independently calculate the embeddings for each view in the target domain and feed the averaged embeddings into the trained GNN for node classification. This approach preserves the core concept of the original non-adaptive methods while leveraging cross-view aggregation for transfer.

Table 1: Experiment on ACM datasets

Methods	ACM1 → ACM2		ACM2 → ACM1	
	ACC	F1	ACC	F1
GCN	0.367	0.253	0.359	0.236
GAT	0.373	0.319	0.363	0.241
AdaGCN	0.431	0.272	0.362	0.386
UDAGCN	0.452	0.283	0.409	0.347
GRADE-N	0.469	0.291	0.343	0.373
ACDNE	0.457	0.261	0.392	0.363
ASN	0.436	0.321	0.437	0.327
SpecReg	0.493	0.308	0.431	0.381
PA	0.506	0.333	0.452	0.377
GraphATA	0.501	0.322	0.395	0.401
HGDA	0.511	0.311	0.441	0.414
SAGA	<u>0.523</u>	<u>0.351</u>	<u>0.454</u>	<u>0.444</u>
CCGA	**0.604**	**0.457**	**0.538**	**0.533**

For adaptive baselines (e.g., AdaGCN (Dai et al., 2022), UDAGCN (Wu et al., 2020), GRADE-N (Wu et al., 2023a), ACDNE (Shen et al., 2020), ASN (Zhang et al., 2021), SpecReg (You et al., 2023), PA (Liu et al., 2024)), we calculate the view-specific embeddings for the source views, and then average the embeddings across all source views to obtain a single representative source embedding for downstream inference in the target domain. Target view embeddings are calculated in target domain according to the formulation, and aligned or regularized via adaptation mechanisms during training as needed. In this experiment, accuracy and Macro-F1 score are utilized as evaluation metrics, since they are standard measures for classification tasks.

5.3 RESULTS AND ANALYSES

The experimental results are summarized in Table 1 and Table 2, where the best results are highlighted in **bold**, and the second-best results are <u>underlined</u>. Some results are directly taken from SAGA (Fang et al., 2026). We have the following findings: Our proposed method boosts the performance of state-of-the-art (SOTA) methods across most evaluation metrics on the ACM and MAG datasets, which proves its superiority over all baseline methods in multi-view graph domain adaptation for node classification tasks. CCGA outperforms other optimal methods, achieving state-of-the-art

Table 2: Experiment on MAG datasets

Methods	Metric	CN→US	US→CN	JP→CN	CN→JP	DE→FR	FR→DE	RU→US	US→RU	DE→CN	CN→DE
GCN	ACC	0.351	0.263	0.327	0.335	0.319	0.302	0.301	0.317	0.281	0.291
	macro-F1	0.217	0.218	0.221	0.201	0.203	0.167	0.233	0.320	0.167	0.223
GAT	ACC	0.325	0.316	0.361	0.309	0.324	0.253	0.265	0.339	0.281	0.316
	macro-F1	0.218	0.320	0.253	0.237	0.212	0.249	0.221	0.323	0.249	0.209
AdaGCN	ACC	0.418	0.483	0.389	0.425	0.431	0.314	0.370	0.435	0.359	0.363
	macro-F1	0.231	0.362	0.257	0.244	0.242	0.285	0.242	0.369	0.265	0.212
UDAGCN	ACC	0.425	0.453	0.451	0.414	0.352	0.364	0.359	0.421	0.356	0.369
	macro-F1	0.243	0.341	0.271	0.275	0.274	0.262	0.239	0.361	0.296	0.253
GRADE-N	ACC	0.421	0.467	0.376	0.404	0.455	0.375	0.353	0.429	0.358	0.362
	macro-F1	0.241	0.352	0.305	0.292	0.286	0.287	0.245	0.347	0.299	0.252
ACDNE	ACC	0.410	0.458	0.422	0.436	0.463	0.359	0.351	0.434	0.360	0.378
	macro-F1	0.225	0.352	0.335	0.275	0.279	0.292	0.231	0.388	0.265	0.291
ASN	ACC	0.416	0.431	0.335	0.432	0.433	0.368	0.289	0.435	0.334	0.372
	macro-F1	0.232	0.387	0.219	0.251	0.248	0.297	0.216	0.226	0.258	0.270
SpecReg	ACC	0.425	0.501	0.363	0.419	0.424	0.387	0.328	0.429	0.330	0.371
	macro-F1	0.247	0.380	0.273	0.263	0.236	0.282	0.213	0.371	0.279	0.303
PA	ACC	0.402	0.529	0.351	0.323	0.416	0.357	0.306	0.317	0.336	0.357
	macro-F1	0.236	0.331	0.295	0.243	0.229	0.299	0.235	0.367	0.286	0.263
GraphATA	ACC	0.407	0.518	0.382	0.341	0.327	0.414	0.351	0.341	0.341	0.321
	macro-F1	0.241	0.339	0.287	0.252	0.291	0.289	0.234	0.351	0.274	0.254
HGDA	ACC	0.417	0.529	0.392	0.343	0.391	0.421	0.361	0.337	0.361	0.317
	macro-F1	0.239	0.333	0.321	0.251	0.284	0.279	0.212	0.344	0.271	0.249
SAGA	ACC	0.553	0.535	0.512	0.557	0.529	0.515	0.377	0.441	0.425	0.381
	macro-F1	0.341	0.426	0.372	0.415	0.426	0.401	0.378	0.397	0.320	0.350
CCGA	ACC	**0.612**	**0.542**	**0.555**	**0.594**	**0.558**	**0.570**	**0.424**	**0.459**	**0.473**	**0.426**
	macro-F1	**0.436**	**0.438**	**0.438**	**0.475**	**0.500**	**0.504**	**0.419**	**0.419**	**0.431**	**0.415**

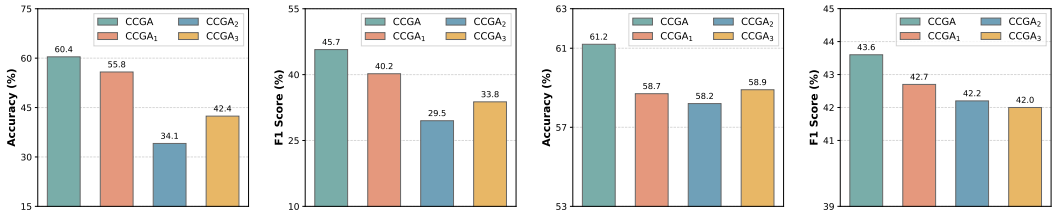
(SOTA) performance across all datasets. Specifically, compared to the strongest baseline SAGA, CCGA achieves an average accuracy improvement of 8.25% and Macro-F1 improvement of 9.75% on the ACM datasets. On the MAG datasets, CCGA achieves an average improvement of 3.88% in accuracy and 6.49% in Macro-F1 score. This demonstrates that discrepancies in class-level structural information exist across different datasets, and our method accurately captures these discrepancies. Consequently, it effectively aligns the structural information of each view across different domains and mitigates class-level structural shifts. Furthermore, the model efficiency experiments (including training time and memory usage) and the t-SNE visualization of learned node embeddings are presented in Appendix D and Appendix E.

5.4 MODEL ABLATION STUDY

To validate the effectiveness of different components in our model, we compare CCGA with three variants on ACM1-ACM2 and CN-US datasets.

- **CCGA**$_1$: CCGA without the intra-domain class separation distance $\tilde{dis}_v^{\,D}$, to show the necessity of separating feature representations of different classes within the same domain.
- **CCGA**$_2$: CCGA without the cross-domain cross-class separation diatance $\tilde{dis}_v^{\,S\leftrightarrow T}$, to show the importance of maintaining the distance between different classes across source domain and target domain.
- **CCGA**$_3$: CCGA without the dynamic confidence weight mechanism τ_{conf} when calculating target domain class centroids, to show the role of filtering low-confidence pseudo-labels in preventing class centroid contamination.

According to Figure 5, we can draw the following conclusions: (1) The experimental results of CCGA consistently outperform all other variants, indicating the rationality of our method. (2) After removing $\tilde{dis}_v^{\,D}$ and $\tilde{dis}_v^{\,S\leftrightarrow T}$, the model performance drops significantly, demonstrating that explicitly maximizing the distance between different classes within and across domains facilitates cross-domain knowledge transfer. (3) After removing the dynamic confidence weight threshold τ_{conf}, the experimental results show a substantial decline, indicating that low-confidence pseudo-labels negatively impact the class centroids. Additional ablation study results across all datasets are detailed in Appendix G.

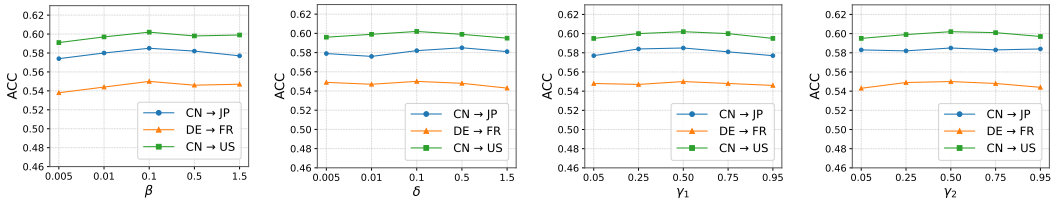


(a) ACM1-ACM2 Acc (b) ACM1-ACM2 F1 (c) CN-US Acc (d) CN-US F1

Figure 5: Model Ablation Study Results.

5.5 HYPERPARAMETER SENSITIVITY ANALYSIS

In this section, we analyze the sensitivity of the parameters of CCGA on DE→FR, CN→US, and CN→JP. As shown in Figure 6(a) - 6(b), the accuracy generally reaches its highest value around $\beta = 0.5$ and $\delta = 0.01$ across the three tasks. This is reasonable since assigning appropriate weights to the source domain classification loss and target domain information entropy encourages the model to extract sufficient class-discriminative features while preserving the confidence of target predictions. When β is set to an extremely low value, the contribution of the source domain classification loss ($\mathcal{L}_S$) is significantly weakened, which causes the model to fail to extract adequate class-discriminative features from the labeled source domain data, consequently leading to a noticeable decline in target node classification performance. Furthermore, each adaptation task may require slightly different suitable values of β and δ because of its inherent graph properties. From Figure 6(c) - 6(d), CCGA maintains competitive performance over a wide range of γ_1 and γ_2, with these two parameters being particularly stable around 0.50 on the three tasks. Only an excessively large or small γ_1 or γ_2 leads to a noticeable decline in performance. As shown in Figure 6, CCGA has compelitive performance on a large range of values, which suggests the stability of CCGA. For a more comprehensive analysis and additional evaluation results of the hyperparameters, please refer to Appendix H.



(a) (b) (c) (d)

Figure 6: Hyperparameter sensitivity analysis of CCGA on different transfer tasks.

6 CONCLUSION

In this paper, we proposed CCGA method to address multi-view graph domain adaptation problem in cross-domain node classification tasks. Through empirical analysis, we reveal the significance of CD-ACD fine-grained multi-view joint alignment method. Our approach leverages CD-ACD to dynamically guide the cross-domain alignment process, characterize and mitigate the structural discrepancies of multi-view graphs at both view level and class level. We evaluate CCGA on the ACM1, ACM2, and MAG datasets to test the method's capability for cross-domain transfer while preserving node class information (Shen et al., 2024a), and these comprehensive experiments validate the superiority of the CCGA method. Furthermore, we will explore the theory of multi-view graph domain adaptation to develop more powerful and generalizable models across various related domains.

7 AI USE STATEMENT

In this work, we used generative AI tools for text polishing, grammar checking, and improving readability. We have not used generative AI tools for conceptual design, theoretical formulation, experimental implementation, or result analysis, and tasks such as AI-driven data generation or literature review are not applicable to this work. We have reviewed all AI-assisted work. Specifically, all AI-edited text was carefully read and verified by the authors to ensure it accurately reflects our original scientific intent, methodology, and conclusions without introducing any misrepresentations. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

Wei Chen, Xingyu Guo, Shuang Li, Zhao Zhang, Yan Zhong, Fuzhen Zhuang, and deqing wang. Learning adaptive distribution alignment with neural characteristic function for graph domain adaptation. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 17271–17294, 2026a.

Wei Chen, Xingyu Guo, Shuang Li, Yan Zhong, Zhao Zhang, Fuzhen Zhuang, Hongrui Liu, Libang Zhang, Guo Ye, and Huimei He. Learning structure-semantic evolution trajectories for graph domain adaptation. *International Conference on Learning Representations*, 2026b.

Quanyu Dai, Xiao-Ming Wu, Jiaren Xiao, Xiao Shen, and Dan Wang. Graph transfer learning via adversarial domain adaptation with graph convolution. *IEEE Transactions on Knowledge and Data Engineering*, 35(5):4908–4922, 2022.

Shaohua Fan, Xiao Wang, Chuan Shi, Emiao Lu, Ken Lin, and Bai Wang. One2multi graph autoencoder for multi-view graph clustering. In *proceedings of the web conference 2020*, pp. 3070–3076, 2020.

Ruiyi Fang, Bingheng Li, Jingyu Zhao, Ruizhi Pu, Qiuhao Zeng, Gezheng Xu, Charles Ling, and Boyu Wang. Homophily enhanced graph domain adaptation. In *Proceedings of the 42nd International Conference on Machine Learning*, ICML'25. JMLR.org, 2025.

Ruiyi Fang, Jingyu Zhao, Shuo Wang, Ruizhi Pu, Bingheng Li, Jiale Cai, Zhihao Li, Zihao Jing, Jian Zhu, Song Tang, et al. Saga: Structural aggregation guided alignment with dynamic view and neighborhood order selection for multiview graph domain adaptation. In *International Conference on Learning Representations*, volume 2026, pp. 95585–95607, 2026.

Jincheng Huang, Yujie Mo, Xiaoshuang Shi, Lei Feng, and Xiaofeng Zhu. Enhancing the influence of labels on unlabeled nodes in graph convolutional networks. *arXiv preprint arXiv:2411.02279*, 2024.

Baoyu Jing, Chanyoung Park, and Hanghang Tong. Hdmi: High-order deep multiplex infomax. In *Proceedings of the web conference 2021*, pp. 2414–2424, 2021.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Chuanyun Lin, Xi Zhou, and Xiao Shen. Hierarchical graph contrastive domain adaptation for multi-source cross-network node classification. *Expert Systems with Applications*, 284:127900, 2025.

Zhiping Lin, Zhao Kang, Lizong Zhang, and Ling Tian. Multi-view attributed graph clustering. *IEEE Transactions on knowledge and data engineering*, 35(2):1872–1880, 2021.

Shikun Liu, Deyu Zou, Han Zhao, and Pan Li. Pairwise alignment improves graph domain adaptation. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 32552–32575. PMLR, 21–27 Jul 2024.

Junyu Luo, Zhiping Xiao, Yifan Wang, Xiao Luo, Jingyang Yuan, Wei Ju, Langechuan Liu, and Ming Zhang. Rank and align: Towards effective source-free graph domain adaptation. In Kate Larson (ed.), *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pp. 4706–4714. International Joint Conferences on Artificial Intelligence Organization, 8 2024. doi: 10.24963/ijcai.2024/520. Main Track.

Junyu Luo, Yuhao Tang, Yiwei Fu, Xiao Luo, Zhizhuo KOU, Zhiping Xiao, Wei Ju, Wentao Zhang, and Ming Zhang. Sparse causal discovery with generative intervention for unsupervised graph domain adaptation. In *Forty-second International Conference on Machine Learning*, 2025.

Junyu Luo, Haoyu Tao, Xiao Luo, Yusheng Zhao, Zhiping Xiao, Dailan He, Wei Ju, Chong Chen, Xian-Sheng Hua, and Ming Zhang. Robust cross supervision with target mining for source-free graph domain adaptation. *IEEE Transactions on Knowledge and Data Engineering*, 2026.

Xin Ma, Yifan Wang, Siyu Yi, Wei Ju, Bei Wu, Ziyue Qiao, Chenwei Tang, and Jiancheng Lv. Pala: Class-imbalanced graph domain adaptation via prototype-anchored learning and alignment. In *IJCAI*, pp. 3198–3207, 2025.

Xin Ma, Yifan Wang, Siyu Yi, Wei Ju, Junyu Luo, Yusheng Zhao, Xiao Luo, and Jiancheng Lv. Dual prototype-enhanced contrastive framework for class-imbalanced graph domain adaptation. *Advances in Neural Information Processing Systems*, 38:48456–48482, 2026.

Yujie Mo, Yuhuan Chen, Yajie Lei, Liang Peng, Xiaoshuang Shi, Changan Yuan, and Xiaofeng Zhu. Multiplex graph representation learning via dual correlation reduction. *IEEE Transactions on Knowledge and Data Engineering*, 35(12):12814–12827, 2023a.

Yujie Mo, Yajie Lei, Jialie Shen, Xiaoshuang Shi, Heng Tao Shen, and Xiaofeng Zhu. Disentangled multiplex graph representation learning. In *International conference on machine learning*, pp. 24983–25005. PMLR, 2023b.

Erlin Pan and Zhao Kang. Multi-view contrastive graph clustering. *Advances in neural information processing systems*, 34:2148–2159, 2021.

Erlin Pan and Zhao Kang. High-order multi-view clustering for generic data. *Information Fusion*, 100:101947, 2023.

Chanyoung Park, Donghyun Kim, Jiawei Han, and Hwanjo Yu. Unsupervised attributed multiplex network embedding. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pp. 5371–5378, 2020.

Liang Peng, Xin Wang, and Xiaofeng Zhu. Unsupervised multiplex graph learning with complementary and consistent information. In *Proceedings of the 31st ACM International Conference on Multimedia*, pp. 454–462, 2023.

Xiaowei Qian, Bingheng Li, and Zhao Kang. Upper bounding barlow twins: A novel filter for multi-relational clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, pp. 14660–14668, 2024.

Ziyue Qiao, Xiao Luo, Meng Xiao, Hao Dong, Yuanchun Zhou, and Hui Xiong. Semi-supervised domain adaptation in graph transfer learning. In Edith Elkind (ed.), *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23*, pp. 2279–2287. International Joint Conferences on Artificial Intelligence Organization, 8 2023. doi: 10.24963/ijcai.2023/253. Main Track.

Ziyue Qiao, Meng Xiao, Weiyu Guo, Xiao Luo, and Hui Xiong. Information filtering and interpolating for semi-supervised graph domain adaptation. *Pattern Recognition*, 153:110498, 2024.

Yalan Qin, Nan Pu, Nicu Sebe, and Guorui Feng. Robust consensus anchor learning for efficient multi-view subspace clustering.

Meng Qu, Jian Tang, Jingbo Shang, Xiang Ren, Ming Zhang, and Jiawei Han. An attention-based collaboration framework for multi-view network representation learning. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pp. 1767–1776, 2017.

Xiao Shen, Quanyu Dai, Fu-lai Chung, Wei Lu, and Kup-Sze Choi. Adversarial deep network embedding for cross-network node classification. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pp. 2991–2999, 2020.

Xiao Shen, Chuanyun Lin, Zhihao Chen, Junwei Gong, and Xi Zhou. Prototypical multi-source graph contrastive domain adaptation. *Pattern Recognition*, pp. 114312, 2026.

Zhixiang Shen, Haolan He, and Zhao Kang. Balanced multi-relational graph clustering. In *Proceedings of the 32nd ACM international conference on multimedia*, pp. 4120–4128, 2024a.

Zhixiang Shen, Shuo Wang, and Zhao Kang. Beyond redundancy: Information-aware unsupervised multiplex graph structure learning. *Advances in neural information processing systems*, 37: 31629–31658, 2024b.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

Kuansan Wang, Zhihong Shen, Chiyuan Huang, Chieh-Han Wu, Yuxiao Dong, and Anshul Kanakia. Microsoft academic graph: When experts are not enough. *Quantitative Science Studies*, 1(1): 396–413, 2020.

Mengzhu Wang, Wenhao Ren, Yu Zhang, Yanlong Fan, Dianxi Shi, Luoxi Jing, and Nan Yin. Gaussian mixture model for graph domain adaptation. In *IJCAI*, pp. 1963–1972, 2025a.

Shuo Wang, Shunyang Huang, Jinghui Yuan, Zhixiang Shen, and Zhao Kang. Cooperation of experts: fusing heterogeneous information with large margin. In *Proceedings of the 42nd International Conference on Machine Learning*, ICML'25. JMLR.org, 2025b.

Hongwei Wen, Can Zhang, Haoyu He, Hanyuan Hang, and Minglong Lei. Progressive graph structure adjustment for homophily shift adaptation. In *Forty-third International Conference on Machine Learning*.

Jun Wu, Jingrui He, and Elizabeth Ainsworth. Non-iid transfer learning on graphs. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pp. 10342–10350, 2023a.

Lingfei Wu, Yu Chen, Kai Shen, Xiaojie Guo, Hanning Gao, Shucheng Li, Jian Pei, and Bo Long. Graph neural networks for natural language processing: A survey. *Foundations and Trends in Machine Learning*, 16(2):119–328, 2023b.

Man Wu, Shirui Pan, Chuan Zhou, Xiaojun Chang, and Xingquan Zhu. Unsupervised domain adaptive graph convolutional networks. In *Proceedings of the web conference 2020*, pp. 1457–1467, 2020.

Liang Yang, Xin Chen, Jiaming Zhuo, Di Jin, Chuan Wang, Xiaochun Cao, Zhen Wang, and Yuanfang Guo. Disentangled graph spectral domain adaptation. In *Forty-second International Conference on Machine Learning*, 2025.

Nan Yin, Mengzhu Wang, Zhenghan Chen, Li Shen, Huan Xiong, Bin Gu, and Xiao Luo. Dream: Dual structured exploration with mixup for open-set graph domain adaption. In *International conference on learning representations*, volume 2024, pp. 19776–19797, 2024.

Nan Yin, Li Shen, Mengzhu Wang, Xinwang Liu, Chong Chen, and Xian-Sheng Hua. Dream: a dual variational framework for unsupervised graph domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.

Yuning You, Tianlong Chen, Zhangyang Wang, and Yang Shen. Graph domain adaptation via theory-grounded spectral regularization. In *... International Conference on Learning Representations*, volume 2023, pp. https–dblp, 2023.

Haodong Zhang, Tao Ren, Changhu Wang, Yifan Wang, Wei Ju, Huaizhi Tang, Junyu Luo, Zimo Wang, Ziyue Qiao, Xian-Sheng Hua, et al. Disco: Diffusion-guided unbiased discriminative learning for unsupervised graph domain adaptation. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1*, pp. 1892–1903, 2026.

Xiaowen Zhang, Yuntao Du, Rongbiao Xie, and Chongjun Wang. Adversarial separation network for cross-network node classification. In *Proceedings of the 30th ACM international conference on information & knowledge management*, pp. 2618–2626, 2021.

Zhen Zhang and Bingsheng He. Aggregate to adapt: Node-centric aggregation for multi-source-free graph domain adaptation. In *Proceedings of the ACM on Web Conference 2025*, pp. 4420–4431, 2025.

Zhen Zhang and Bingsheng He. Towards unsupervised open-set graph domain adaptation via dual reprogramming. *Advances in Neural Information Processing Systems*, 38:53320–53343, 2026.

Zhen Zhang, Meihan Liu, Anhui Wang, Hongyang Chen, Zhao Li, Jiajun Bu, and Bingsheng He. Collaborate to adapt: Source-free graph domain adaptation via bi-directional adaptation. In *Proceedings of the ACM Web Conference 2024*, pp. 664–675, 2024.

Jiaming Zhuo, Ziyi Ma, Yintong Lu, Yuwei Liu, Kun Fu, Di Jin, Chuan Wang, Wu Wenning, Zhen Wang, Xiaochun Cao, et al. A closer look at graph transformers: Cross-aggregation and beyond. *Advances in Neural Information Processing Systems*, 38:136781–136807, 2026.

A DATASET EXPERIMENT DETAILS

The ACM1 and ACM2 graphs are labeled into three classes: database, wireless communication, and data mining. As the subset of OGBN-MAG (Wang et al., 2020), MAG is partitioned into six subgraphs (CN, US, DE, FR, JP, and RU), each corresponding to a country. From these country subgraphs, we extract the ten most frequent classes. Any two of these country subgraphs can be used as the source and target datasets in the multi-view graph node classification experiment. Details of the experimental datasets' statistics are provided in Table 1.

Table 1: Statistics of experimental datasets

Types	Datasets	#Node	#Relation	#Edge	#Feat	#Label
ACM	ACM1	3,025	PAP PSP	26,256 2,207,736	512	3
	ACM2	4,019	PAP PSP	57,853 4,338,213	512	3
MAG	CN	25,467	PAP PP	157,303 43,683	128	10
	US	55,392	PAP PP	12,332,742 289,833	128	10
	JP	12,651	PAP PP	41,165 20,747	128	10
	FR	12,052	PAP PP	51,739 30,915	128	10
	DE	17,919	PAP PP	77,100 45,528	128	10
	RU	3,690	PAP PP	3,315 3,539	128	10

B BASELINES

In this work, we compare a series of representative multi-view graph domain adaptation baseline methods for node classification. We start with non-adaptive methods, where GCN (Kipf & Welling, 2016) serves as a single-view baseline, performing standard graph convolutions without any domain adaptation. GAT (Veličković et al., 2017) utilizes expressive neighborhood aggregation to provide an attention-based non-adaptive alternative. To address complex graph structures, AdaGCN (Dai et al., 2022) proposes an adaptive propagation method that tailors message passing to node-level or edge-level cues, thereby responding to diverse neighborhood structures in a multi-view environment; PA (Liu et al., 2024) addresses structural shift by recalibrating edge influences; UDAGCN (Wu et al., 2020) represents a dual graph convolutional network learning framework for unsupervised GDA, capturing knowledge at local and global levels and utilizing adversarial training to align cross-view representations.

ACDNE (Shen et al., 2020) is an adaptation-oriented method that aligns cross-view targets and fuses representations across views. ASN (Zhang et al., 2021) leverages domain-specific features within networks to extract domain-invariant shared representations across multiple networks. In measuring distribution shift, GRADE-N (Wu et al., 2023a) introduces graph subtree discrepancy to quantify the distribution shift between the source graph and the target graph; SpecReg (You et al., 2023) achieves performance improvements through regularization inspired by optimal transport domain adaptation and graph filter theory; GraphATA (Zhang & He, 2025) performs node-wise personalized aggregation of source GNNs for multi-source source-free unsupervised graph domain adaptation. HGDA (Fang et al., 2025) is a framework that mitigates cross-domain discrepancies by explicitly aligning homophilic, heterophilic, and attribute signals through specialized filters, thereby improving node classification across source and target graphs. SAGA (Fang et al., 2026) utilizes structural aggregation distance to evaluate structural discrepancies across different hops and views, aligning multi-view graph data through dynamic view and neighbor order selection.

Table 2: Comparison of Training Time, Memory Usage, and Accuracy on ACM and MAG dataset.

Dataset	Method	Training Time (Normalized w.r.t. CCGA)	Memory Usage (Normalized w.r.t. CCGA)	Accuracy(%)
ACM1 → ACM2	UDAGCN	0.836	1.055	45.2
	ASN	1.089	1.486	43.6
	PA	0.984	1.197	50.6
	HGDA	0.768	1.137	51.1
	SAGA	0.960	1.184	52.3
	CCGA	1	1	60.4
ACM2 → ACM1	UDAGCN	0.873	0.894	40.9
	ASN	1.156	1.248	43.7
	PA	1.038	1.039	37.7
	HGDA	0.817	1.012	44.1
	SAGA	1.026	1.005	45.4
	CCGA	1	1	53.8
CN → US	UDAGCN	0.948	1.224	42.5
	ASN	1.343	1.937	41.6
	PA	0.948	1.417	40.2
	HGDA	0.959	1.508	41.7
	SAGA	1.062	1.420	55.3
	CCGA	1	1	61.2
US → CN	UDAGCN	0.928	1.294	45.3
	ASN	1.319	1.947	43.1
	PA	0.928	1.407	52.9
	HGDA	0.949	1.508	52.9
	SAGA	1.051	1.420	53.5
	CCGA	1	1	54.2
JP → CN	UDAGCN	1.012	1.398	45.1
	ASN	1.415	2.107	33.5
	PA	1.013	1.544	35.1
	HGDA	1.016	1.659	39.2
	SAGA	1.122	1.546	51.2
	CCGA	1	1	55.5
CN → JP	UDAGCN	0.810	1.378	41.4
	ASN	1.057	2.127	43.2
	PA	0.833	1.572	32.3
	HGDA	0.844	1.648	34.3
	SAGA	0.932	1.546	55.7
	CCGA	1	1	59.4
DE → FR	UDAGCN	0.926	1.361	35.2
	ASN	1.241	2.137	43.3
	PA	0.937	1.566	41.6
	HGDA	0.939	1.563	39.1
	SAGA	1.060	1.571	52.9
	CCGA	1	1	55.8
FR → DE	UDAGCN	0.907	1.428	36.4
	ASN	1.185	2.168	36.8
	PA	0.943	1.544	35.7
	HGDA	0.994	1.561	42.1
	SAGA	1.083	1.571	51.5
	CCGA	1	1	57.0
RU → US	UDAGCN	0.968	1.461	35.9
	ASN	1.243	2.108	28.9
	PA	0.975	1.612	30.6
	HGDA	0.983	1.542	36.1
	SAGA	1.085	1.589	37.7
	CCGA	1	1	42.4
US → RU	UDAGCN	0.929	1.468	42.1
	ASN	1.263	2.092	43.5
	PA	0.9431	1.526	31.7
	HGDA	0.963	1.539	33.7
	SAGA	1.054	1.589	44.1
	CCGA	1	1	45.9
CN → DE	UDAGCN	0.904	1.336	36.9
	ASN	1.187	2.007	37.2
	PA	0.957	1.546	35.7
	HGDA	0.969	1.643	31.7
	SAGA	1.065	1.529	38.1
	CCGA	1	1	42.6
DE → CN	UDAGCN	0.931	1.403	35.6
	ASN	1.228	2.119	33.4
	PA	0.937	1.540	33.6
	HGDA	0.936	1.642	36.1
	SAGA	1.027	1.529	42.5
	CCGA	1	1	47.3

C EXPERIMENTAL SETUP

The experiments are implemented on the PyTorch platform, utilizing an Intel(R) Xeon(R) Gold 6248R CPU @ 3.00GHz and an NVIDIA GeForce RTX 3090 24GB GPU. We construct a multi-view graph propagation module alongside a two-layer MLP encoder, and the model is trained using the Adam optimizer with a learning rate ranging from 0.0005 to 0.001. To prevent over-fitting, the dropout rate is set to 0.1. Furthermore, the weight decay coefficient is set to $\in \{1\text{e-}4, \cdots, 5\text{e-}3\}$. The model adopts a two-stage representation learning architecture, where the hidden dimension is set to 256, the embedding dimension is set to 128, and the neighborhood propagation order is set to 2. For the different transfer tasks, the batch size is set to 1000 for ACM1-ACM2 and 10000 for the MAG-based tasks. Under the same dataset partitioning, we repeatedly train and test our model five times, reporting the node classification performance on both the source and target domains, including ACC and Macro-F1.

D MODEL EFFICIENT EXPERIMENT

Model Complexity: Here, we analyze the computational complexity of the proposed CCGA. The computational complexity primarily depends on the filter layers and key loss function. For a given graph G, let N represent the total number of nodes in the graph, V represent view number and d the feature dimension. For Scalable Graph Encoding f_Θ: Includes matrix multiplication with $O(N^2 \cdot d)$. CCGA per-iteration complexity is dominated by multi-view propagation $O(VN \cdot d)$, and a forward-only CD-ACD computation $O(VN^2 d)$. Combining these yields an overall complexity of $O(VN \cdot d + VN^2 \cdot d)$, where the quadratic CD-ACD term does not incur gradient cost, making the practical training overhead similar to existing GDA and MGDA methods. Thus, the total complexity of CCGA is still: $O(VN^2 \cdot d)$. Note that HGDA is inherently a single-view method. If we extend HGDA to the multi-view MGDA setting, it would inherit the same computational overhead, resulting in a complexity of $O(VN^2 \cdot d)$. Therefore, although CCGA incorporates class centroid computation CD-ACD, its computational overhead remains comparable to that of HGDA and the existing MGDA method SAGA; these three methods do not differ by orders of magnitude in asymptotic complexity.

Model Efficient Experiment: To further investigate the efficiency of CCGA, we compared the training time and GPU memory usage of common baselines SAGA. As shown in Table 2, the evaluation results on ACM and MAG dataset further demonstrate that our method achieves superior performance with tolerable computational and storage overhead.

E VISUALIZATION

To evaluate the quality of learned representations, we visualize the node embeddings generated by CCGA using t-SNE on both ACM1 and ACM2 datasets. We compare three different configurations: (a) embeddings using only the PAP structural view, (b) embeddings using only the PSP structural view, and (c) embeddings incorporating both PAP and PSP views .

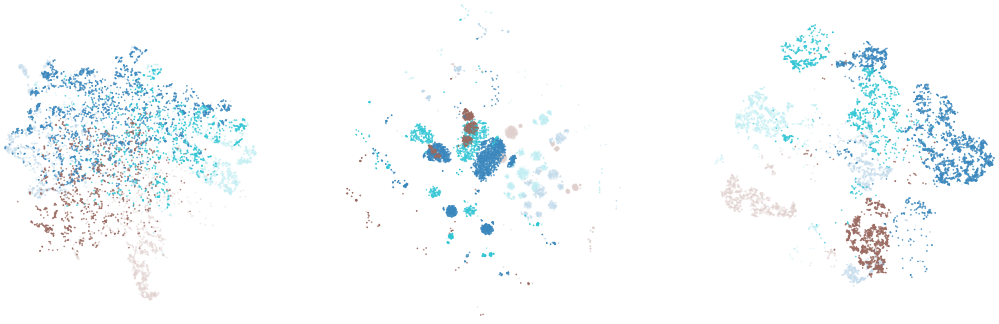
As shown in Figure 1, each color represents a distinct node class, while within each class, lighter and darker shades denote nodes from the source and target domains, respectively. All three configurations achieve clear separation between different classes, demonstrating that CCGA effectively preserves inter-class discriminability.

However, significant differences emerge in the alignment of cross-domain nodes within the same class. In subfigures (a) and (b), where only single views are used, source and target domain nodes of the same class exhibit noticeable distribution gaps. In contrast, subfigure (c) shows substantially improved overlap between domains within each class, indicating that the joint representation better aligns the source and target distributions.

These results suggest that integrating multiple structural views provides complementary information that enhances domain adaptation performance. The multi-view approach not only maintains class discrimination but also facilitates better cross-domain alignment, ultimately yielding more transferable node representations for classification tasks.

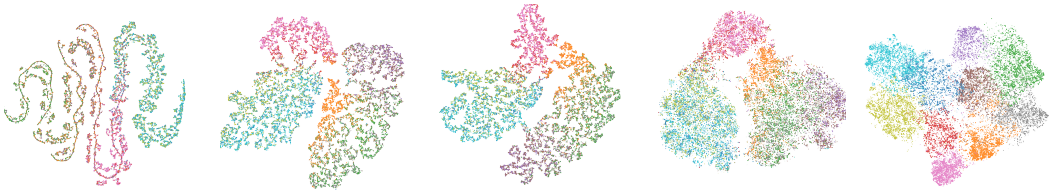
To further demonstrate the advantage of our proposed method, we also visualize the embedding results on MAG generated by competitive methods SAGA, UDAGCN, PA, and ASN, which are

shown in Figure 2. Our CCGA method produces a compact alignment structure. In other words, our method has the best domain distance between source target and the most distinct boundaries between different classes.



(a) only PAP (b) only PSP (c) both PAP and PSP

Figure 1: t-SNE visualization of node embeddings on MAG datasets. (a) Embeddings using only the PAP view; (b) Embeddings using only the PSP view; (c) Embeddings using the joint PAP and PSP views. Each color represents a distinct class, with lighter and darker shades indicating source and target domain nodes within the same class.



(a) UDAGCN (b) PA (c) ASN (d) SAGA (e) CCGA

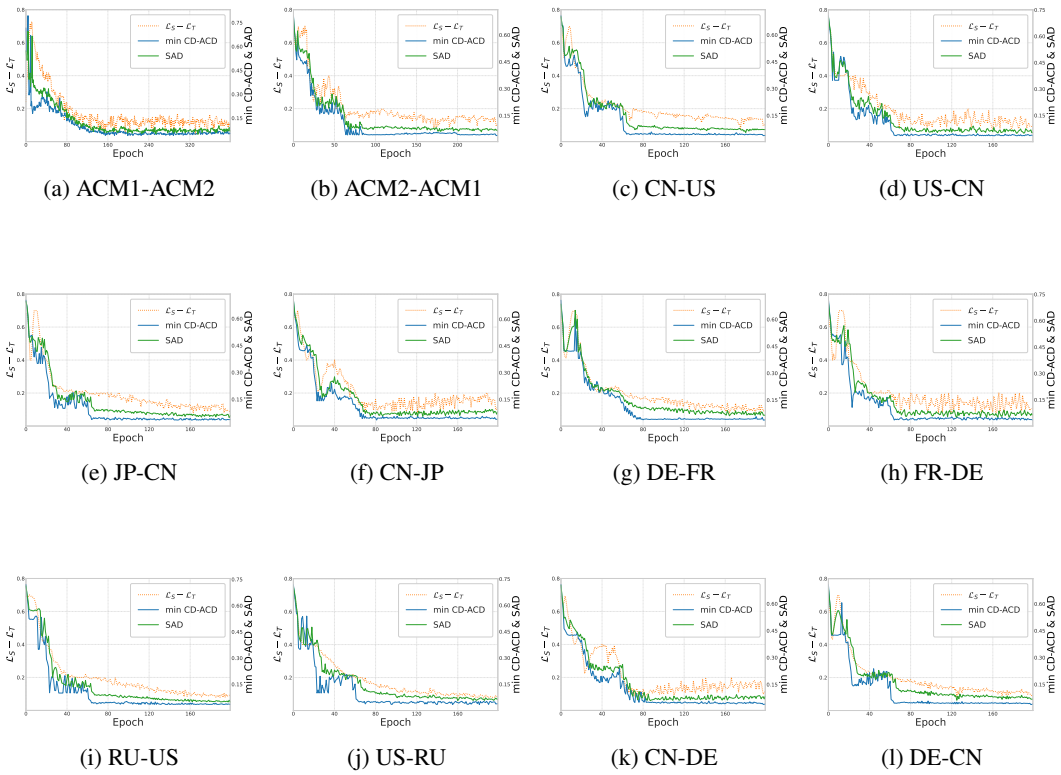
Figure 2: t-SNE visualization of node embeddings on MAG datasets.

F EXPERIMENT DETAIL OF EMPIRICAL STUDY

We evaluate the proposed CCGA method under a unified multi-view graph domain adaptation protocol on both the ACM and MAG benchmarks. To quantify the performance gap between source domain and target domain, we calculate $\mathcal{L}_S - \mathcal{L}_T$. Specifically, $\mathcal{L}_S$ follows the definition in Equation 10, where $\mathcal{L}_T$ does not correspond to the target-domain classification loss in Equation 1. The latter is defined as:

$$\mathcal{L}_T = -\frac{1}{N^T} \sum_{j=1}^{N^T} \sum_{c=1}^{C} P_{j,c}^T \log \left(\frac{\exp(O_{j,c}^T)}{\sum_{k=1}^{C} \exp(O_{j,k}^T)} \right). \tag{1}$$

In this context, the difference $\mathcal{L}_S - \mathcal{L}_T$ serves as a measure of the relative performance between source domain and target domain. We conducted an empirical study across all datasets. As illustrated in Figure 3, it can be observed that as the training epochs increase, the cross-domain classification loss discrepancy ($\mathcal{L}_S - \mathcal{L}_T$) represented by the left y-axis progressively decreases, the value of min CD-ACD from our CCGA method and SAD metric from SAGA method decreases correspondingly, min CD-ACD converges to a lower discrepancy value than SAD. Based on this observation, we can conclude that a smaller CD-ACD indicates that the source and target domains achieve higher-quality cross-domain alignment at the class level, and CD-ACD is more effective than the existing SAD metric.



(a) ACM1-ACM2 (b) ACM2-ACM1 (c) CN-US (d) US-CN

(e) JP-CN (f) CN-JP (g) DE-FR (h) FR-DE

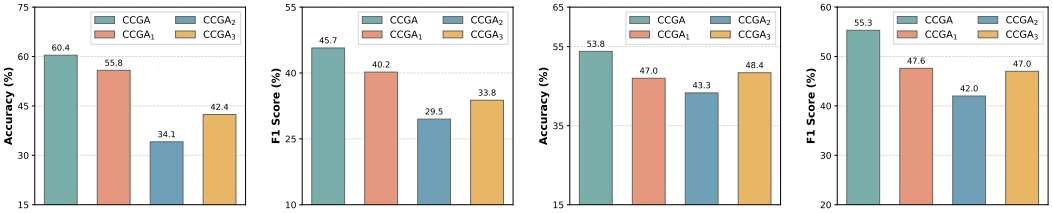
(i) RU-US (j) US-RU (k) CN-DE (l) DE-CN

Figure 3: Training dynamics curves of CD-ACD on all dataset.

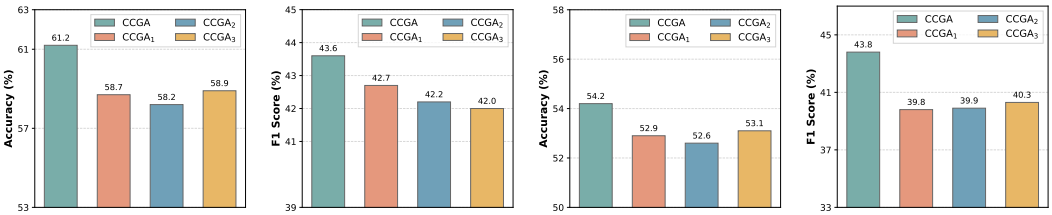
G MODEL ABLATION STUDY

In this section, we further analyze the effectiveness of the various components of our proposed CCGA method on the ACM and MAG datasets. To systematically evaluate the contribution of each module, we compare the full CCGA model with its designed variants. As shown in the subfigures of Figure 4, the experimental results of CCGA consistently outperform all other variants, indicating the rationality of our model. After removing $\tilde{dis}_v^{\,D}$ and $\tilde{dis}_v^{\,S\leftrightarrow T}$ component, the model performance drops significantly, demonstrating that explicitly maximizing the distance between different classes within and across domains facilitates cross-domain knowledge transfer. After removing the dynamic confidence weight threshold τ_{conf} component, the experimental results show a substantial decline, indicating that low-confidence pseudo-labels negatively impact the class centroids.

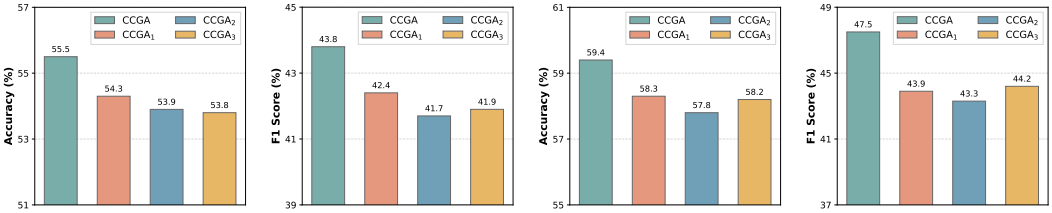
Generally, removing the intra-domain class separation distance $\tilde{dis}_v^{\,D}$ results in a relatively minor performance decline. This is because, even without the explicit intra-domain distance maximization constraint, the model is still optimized by the source domain classification loss and the target domain information entropy optimization. Removing the dynamic confidence weight threshold τ_{conf} leads to a more obvious performance drop, as the absence of this threshold allows low-confidence incorrect pseudo-labels to directly participate in the computation, thereby affecting the construction of target domain class centroids. Removing the cross-domain cross-class separation distance $\tilde{dis}_v^{\,S\leftrightarrow T}$ causes the most significant performance degradation. In unsupervised graph domain adaptation tasks, without reasonably distinguishing a specific class in source domain from other classes in target domain, the model is prone to erroneously aligning class A in source domain with class B in target domain during the cross-domain alignment process.



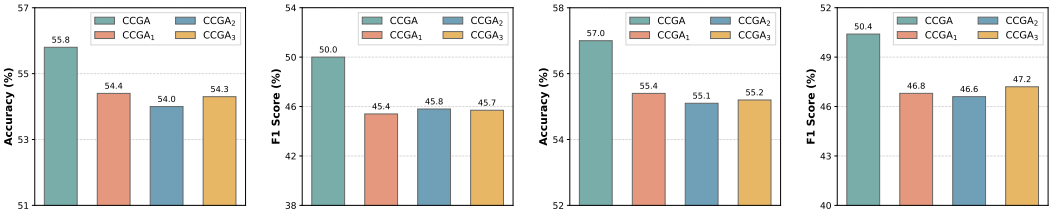
(a) ACM1-ACM2 Acc (b) ACM1-ACM2 F1 (c) ACM2-ACM1 Acc (d) ACM2-ACM1 F1



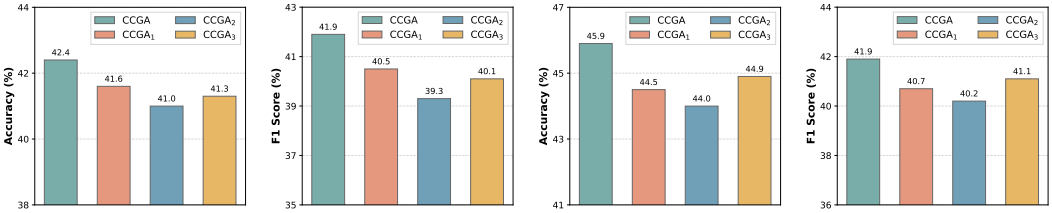
(e) CN-US Acc (f) CN-US F1 (g) US-CN Acc (h) US-CN F1



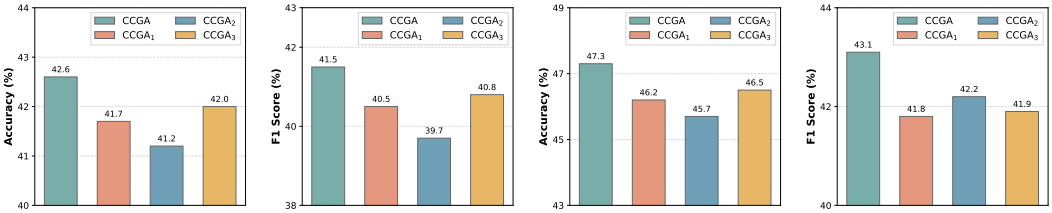
(i) JP-CN Acc (j) JP-CN F1 (k) CN-JP Acc (l) CN-JP F1



(m) DE-FR Acc (n) DE-FR F1 (o) FR-DE Acc (p) FR-DE F1



(q) RU-US Acc (r) RU-US F1 (s) US-RU Acc (t) US-RU F1



(u) CN-DE Acc (v) CN-DE F1 (w) DE-CN Acc (x) DE-CN F1

Figure 4: Model Ablation Study Results across all datasets.

H PARAMETER ANALYSIS

The effectiveness of Graph Domain Adaptation is influenced by the hyperparameters β, δ, γ_1, γ_2, and α, which provide flexibility for adjusting the relative importance of different loss terms. In our experiments, β and δ are chosen from the set $\{0.005, 0.01, 0.1, 0.5, 1.5\}$. The parameters γ_1 and γ_2 are selected from the range $(0, 1)$. The graph propagation hyperparameter α is chosen from the set $\{0.005, 0.01, 0.05, 0.1, 0.5\}$.

ACM Dataset: The ACM dataset often exhibits diverse structural patterns across different distributions, with varying levels of similarity between source and target graphs. These diverse structural characteristics necessitate a careful balance between domain alignment and label information from the source graph. Therefore, β is selected from $\{0.1, 0.5\}$ to balance the source classification loss and target predictions. For γ_1 and γ_2, values are selected from $\{0.4, 0.5\}$ to maintain class separability. A small graph propagation hyperparameter α is selected from $\{0.005, 0.01\}$ to prevent over-smoothing and capture relevant local structures without introducing excessive noise from distant nodes.

MAG Dataset: The MAG dataset is large and diverse, containing losts of classes with various relationships and rich metadata.Structural and attribute alignment are key factors. In this context. β and δ are selected from $\{0.01, 0.1\}$ to balance the classification supervision with alignment constraints. Due to the complex inter-class relationships, γ_1 and γ_2 are selected from $\{0.4, 0.5\}$ to enforce cross-class separation. To capture local and global structural information, α is selected from $\{0.05, 0.1\}$, enabling propagation while reducing noise.

In this section, we analyze the sensitivity of the parameters of our method on the ACM dataset and MAG dataset. As shown in the subfigures of Figure 5, when β is set to an extremely low value, the contribution of the source domain classification loss is weakened, which causes the model to fail to extract adequate class-discriminative features from the labeled source domain data, leading to a decline in target node classification performance. For γ_1 and γ_2, extremely small values cannot effectively maximize the distance between different classes, while extremely large values will influence the cross-domain alignment structure. Furthermore, an extremely large α could prevent the aggregation of neighborhood topological information.

Table 3: Experiment hyperparameter setting Value.

Types	Datasets	β	δ	γ_1	γ_2	α
ACM	ACM1 → ACM2	0.5	0.01	0.5	0.5	0.01
	ACM2 → ACM1	0.5	0.01	0.35	0.4	0.01
MAG	CN → US	0.1	0.1	0.5	0.5	0.05
	US → CN	0.1	0.1	0.5	0.5	0.05
	JP → CN	0.5	0.1	0.5	0.5	0.01
	CN → JP	0.1	0.5	0.4	0.4	0.05
	DE → FR	0.1	0.1	0.5	0.5	0.05
	FR → DE	0.1	0.01	0.3	0.3	0.01
	RU → US	0.1	0.1	0.4	0.4	0.05
	US → RU	0.1	0.01	0.4	0.4	0.05
	DE → CN	0.1	0.1	0.5	0.4	0.05
	CN → DE	0.01	0.1	0.4	0.5	0.1

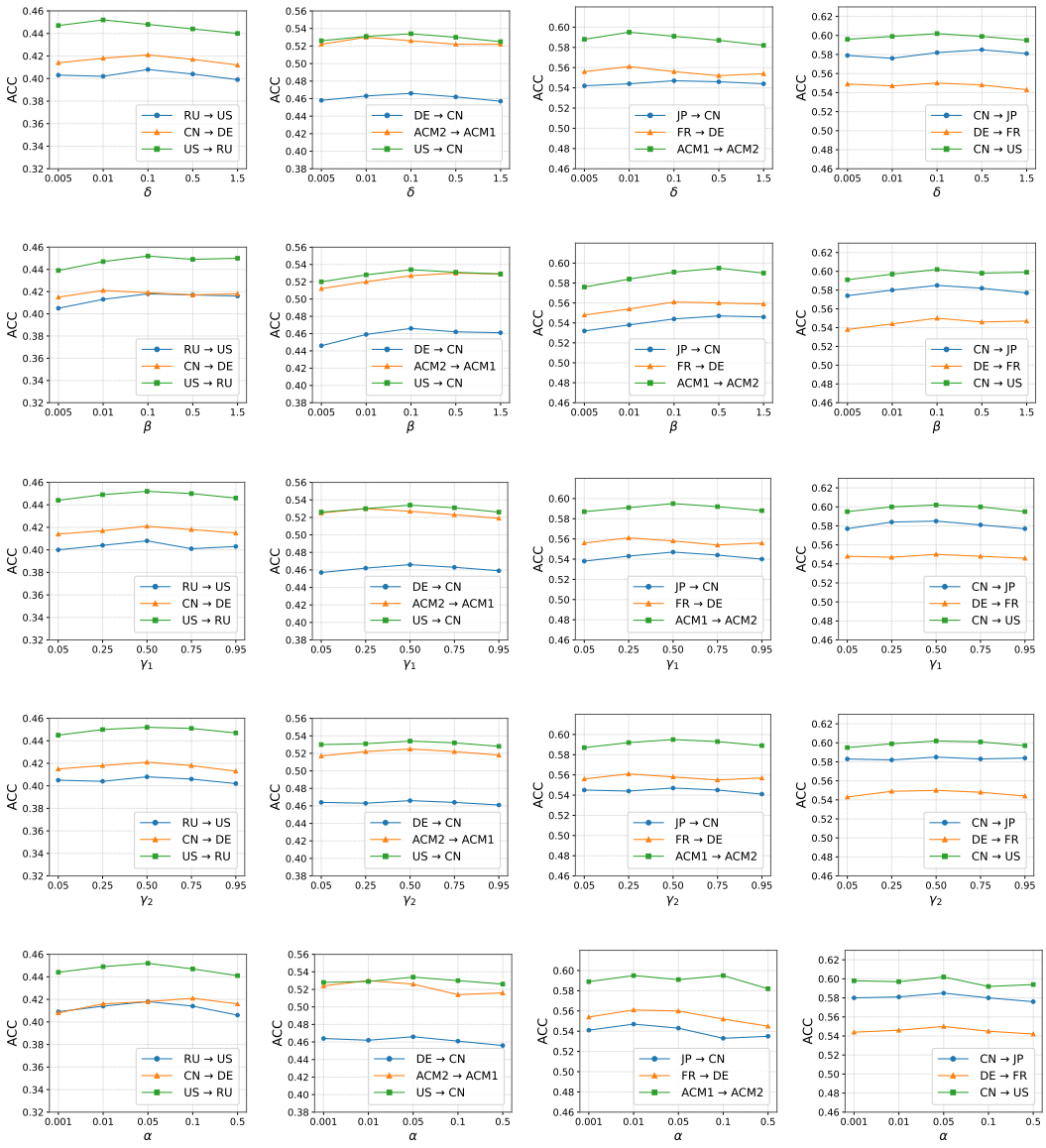


Figure 5: The influence of hyper-parameters β, δ, γ_1, γ_2, and α on ACM and MAG datasets.

I DESCRIPTION OF ALGORITHM CCGA

Algorithm 1: The proposed algorithm CCGA

Input:
Source node feature matrix X^S; source multi-view graph adjacency matrices $\{A_v^S\}_{v=1}^V$; Target node feature matrix X^T; Target multi-view graph adjacency matrices $\{A_v^T\}_{v=1}^V$; source node label matrix Y^S; maximum number of iterations n

for $it = 1$ **to** n **do**
- $Z_v^S = f_\Theta(X_{v,(K)}^S)$
- $Z_v^T = f_\Theta(X_{v,(K)}^T)$ // embedding of source and target graph
- $\hat{X}_v^S = g_\Theta(Z_v^S)$
- $\hat{X}_v^T = g_\Theta(Z_v^T)$ // decoding of source and target graph to alleviate representation collapse
- $O^D = h_\Phi(Z_U^D)$ // global multi-view information mapping
- $\bar{z}_{m,v}^S$ and $\bar{z}_{m,v}^T$ constructed by $Y_{i,m}^S$ and $W_{j,m}^T$
- $CD\text{-}ACD_v$ is calculated once per epoch by Formula 3 algorithm
- Z_k^S and Z_k^T through cross-domain alignment by Formula 8 algorithm.
- Z_v^S and Z_v^T through intra-domain alignment by Formula 9 algorithm.
- Domain adaptive between S and T // adaptive guided by $CD\text{-}ACD_v$
- $O_{i,c}^S$ constrained by $Y_{i,c}^S$ and $O_{j,c}^T$ constrained by information entropy
- Calculate the overall loss with Equation 12
- Update all parameters of the framework according to the overall loss

end

Predict the labels of target graph nodes based on the trained framework.

Output: Classification result $\hat{Y}^T$