From Semantics to Spectrum: A New Lens on Graph Augmentation Strategy

Xiangping Zheng[1,2,3,4*], Xiuxin Hao[1,2*], Bo Wu[5], Wei Li[1,4†], Bin Ren[2,6], Bin Tang[1,2], Yuhui Guo[7], Xun Liang[8], Zhiwen Yu[1,4]

[1]College of Computer Science and Technology, Harbin Engineering University, Harbin, Heilongjiang, China
[2]Qingdao Innovation and Development Center of Harbin Engineering University, Qingdao, Shandong, China
[3]Sanya Nanhai Innovation and Development Base of Harbin Engineering University, Sanya, Hainan, China
[4]Modeling and Emulation in E-Government National Engineering Laboratory, Harbin, Heilongjiang, China
[5]Xiangjiang Laboratory, Changsha, Hunan, China
[6]College of Physics and Optoelectronic Engineering, Harbin Engineering University, Harbin, Heilongjiang, China
[7]Beijing Jinghang Research Institute of Computing and Communication, Beijing, China
[8]School of Information, Renmin University of China, Beijing, China
{xpzheng, hxx429292}@hrbeu.edu.cn

Abstract

Graph augmentation is a cornerstone of effective graph contrastive learning, yet existing methods often rely on random designed perturbations, which may distort latent semantics and impair representation quality. In this work, we argue that semantic consistency can be effectively approximated by low-frequency components in the spectral domain, offering a principled proxy for guiding augmentation. Based on this insight, we propose Frequency-Aware Graph Contrastive Learning (**FA-GCL**), a novel framework that explicitly preserves low-frequency signals while selectively perturbing high-frequency components. By aligning augmentation with frequency-aware decomposition, FA-GCL generates diverse yet semantically coherent views, mitigating semantic drift and enhancing representational discrimination. Extensive experiments across multiple benchmarks demonstrate that FA-GCL consistently outperforms state-of-the-art baselines with statistically significant gains, validating its exclusive merits.

Introduction

Graph contrastive learning (GCL) has emerged as a powerful self-supervised paradigm for learning robust and discriminative graph representations by contrasting multiple augmented views of the same graph (Xu et al. 2025; Wei et al. 2023; Yan et al. 2024; Zheng et al. 2024a). As a foundational technique, GCL enables models to capture invariant and semantically meaningful structural patterns, leading to impressive performance across diverse graph learning tasks and domains (Jiang et al. 2025; Zhang et al. 2024; Zheng et al. 2024b).

A key driver of GCL's success lies in designing augmentations that introduce sufficient variability while preserving the semantic integrity of the input graph. To achieve semantic consistency, recent studies (Li et al. 2022; Ji et al. 2024; Yan et al. 2024) focus on identifying and retaining core semantic components. However, widely used rule-based or random perturbations offer limited control and can inadvertently alter functionally critical substructures. For instance, in molecular graphs, even minor modifications to key functional groups can drastically compromise downstream predictions (Li et al. 2022). This reveals a fundamental tension: while semantic invariance is desired, the heuristics employed to approximate it may paradoxically disrupt the very structures that encode it. Hence, an ideal augmentation strategy should not only maintain semantic consistency but also preserve structurally meaningful subgraphs (Wei et al. 2023; Zheng et al. 2022). Yet achieving reliable semantic preservation remains challenging, especially in complex graphs where semantics are implicit, entangled, and difficult to model directly. To address this, we propose a paradigm shift: *instead of relying on explicit structural heuristics, we approximate semantic consistency via low-frequency invariance in the graph spectral domain*, which naturally filters out noise and enhances robustness against semantically disruptive augmentations.

Spectral graph analysis provides a principled framework for disentangling global and local structural patterns (Lin, Chen, and Wang 2023; Zheng et al. 2025). As illustrated in **Fig. 1(a)**, a graph signal can be decomposed into low- and high-frequency components, where the heatmap indicates the magnitude of spectral coefficients. Low-frequency components vary smoothly across the graph, encoding global semantics such as clusters or communities, while high-frequency components capture localized variations and noise (Xiao et al. 2023; Qin et al. 2024). To evaluate the semantic contribution of different frequency bands, we design a controlled experiment: given a graph $\mathcal{G}$, we generate two positive views by injecting Gaussian noise into (1) a subset of low-frequency components and (2) a subset of high-frequency components. As shown in **Fig. 1(b)**, perturbing low-frequency components leads to a substantial drop in performance, whereas perturbing high-frequency components often yields some improvement. This observation further indicates that low-frequency signals act as robust carriers of semantic information, and preserving them serves as a reliable proxy for maintaining semantic consistency in contrastive graph learning.

*These authors contributed equally.
†Corresponding author.

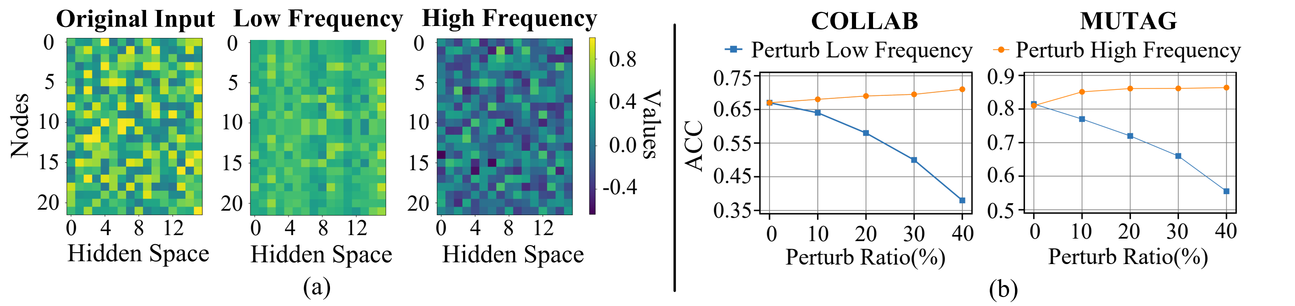

Figure 1: (a) Visualization of raw, low-frequency, and high-frequency features in the spectral domain on the IMDB-BINARY dataset. (b) Evaluation of perturbations applied to different frequency components under varying perturbation ratios on the COLLAB and MUTAG datasets.

Building on this insight, we propose **Frequency-Aware Graph Contrastive Learning (FA-GCL)**, a framework that leverages frequency-domain decomposition to guide graph augmentation. FA-GCL adaptively perturbs high-frequency components to enrich structural diversity, while preserving low-frequency components to maintain semantic alignment, producing contrastive views that are both diverse and semantically consistent. Furthermore, conventional GCL methods often rely on static or random strategies for negative sample construction, overlooking the evolving nature of learned representations. To address this, we design a *Dynamic Difficulty Adjustment (DDA)* module that gradually reduces perturbations on low-frequency components, generating increasingly challenging negatives in sync with the model's learning progress. This DDA module enhances contrastive optimization, sharpens decision boundaries, and ultimately yields more discriminative and robust graph representations. In summary, our main contributions are as follows:

- We propose FA-GCL, a framework designed to enhance the representation quality of graph embeddings. We consider the low-frequency invariance of graphs as a proxy for semantic consistency, thereby facilitating frequency-aware node augmented representation.
- We designed a dynamic difficulty adjustment mechanism to generate negative samples, which gradually reduces the disturbance on low-frequency components and produces negative samples with increasing difficulty that matches the model learning stage.
- Extensive experiments on multiple real-world benchmarks demonstrate that FA-GCL consistently outperforms state-of-the-art baselines with statistically significant margins.

Related Work

Graph Contrastive Learning Graph contrastive learning (GCL) has emerged as a powerful self-supervised paradigm across domains such as social networks (Shen et al. 2023) and bioinformatics (Peng et al. 2024). Early methods, including GRACE (Zhu et al. 2020) and GraphCL (You et al. 2020), rely on random augmentations (e.g., edge deletion, feature masking) that often overlook structural semantics (Chang et al. 2021; Li et al. 2022). Recent advances have explored rule-based augmentations—such as Dual-Prism (Xia et al. 2025), CI-GCL (Tan et al. 2024), DRGCL (Ji et al. 2024), DGPM (Yan et al. 2024), GCS (Wei et al. 2023), RGCL (Li et al. 2022), and AD-GCL (Suresh et al. 2021)—yet they still lack fine-grained control over perturbation frequency and intensity, which is crucial for preserving semantic integrity. To address this gap, we propose FA-GCL, a frequency-aware graph contrastive learning framework that selectively perturbs high-frequency components while preserving low-frequency semantics. This design not only enhances the discriminative power of learned representations but also improves their generalization across diverse graph tasks.

Spectral Graph Analysis Spectral analysis provides a principled framework for capturing multi-scale structural patterns in graphs by decomposing signals into distinct frequency components. It has shown strong utility in tasks such as anomaly detection (Cao et al. 2021), cross-domain alignment (Xiao et al. 2023), and collaborative filtering (Qin et al. 2024). Motivated by these insights, we introduce a **frequency-adaptive graph augmentation** scheme that tailors perturbations according to spectral characteristics. Unlike random or rule-based approaches, our method preserves low-frequency components that encode semantic information, while adaptively perturbing high-frequency components that capture fine-grained structural details. This spectral perspective produces contrastive views that are informative, offering a new paradigm for GCL in the frequency domain.

Methodology

This work presents a frequency-aware framework to improve graph contrastive learning by explicitly modeling spectral properties of graph signals. As illustrated in **Fig. 2**, our approach comprises two key components: **Frequency-Aware Computation** and **Frequency-Driven Sample Generation**. The former decomposes the input graph into low- and high-frequency structures based on spectral smoothness, while the latter utilizes this decomposition to construct informative contrastive views. Central to our design is the insight that *low-frequency structural consistency can act as a robust surrogate for semantic consistency*, providing a principled mechanism to preserve task-relevant information during augmentation.

Frequency-Aware Computation

Conventional graph augmentation methods—whether random, rule-based, or semantic-aware—often may fail to retain

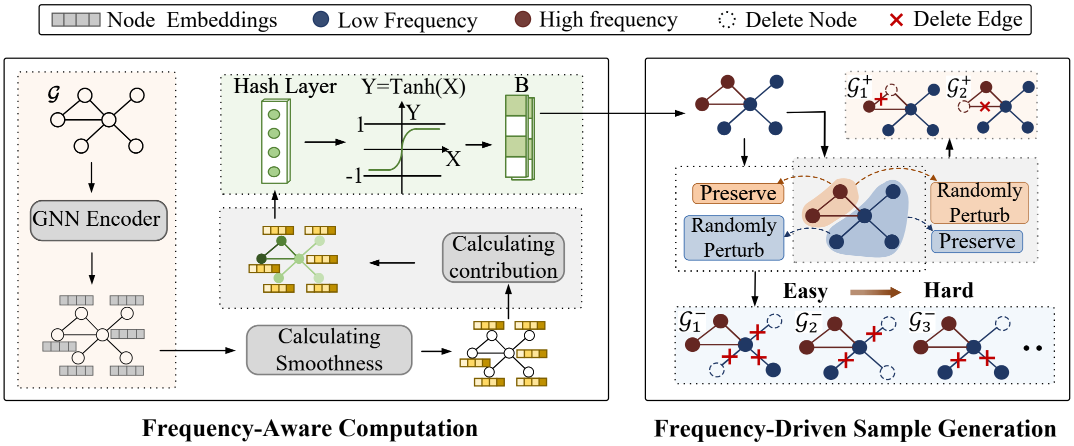

Figure 2: The architecture of our approach consists of a frequency-aware computation module and a frequency-driven sample generation module, working in tandem to enhance graph representation learning.

meaningful structures when semantics are implicit or entangled. We address this by drawing on spectral graph theory to approximate semantic content via frequency decomposition. Prior works (Tan et al. 2024; Cao et al. 2021) show that **low-frequency components** capture smooth, global structures (e.g., community-level regularities), which tend to correlate with semantic identity. In contrast, **high-frequency components** encode localized, high-variance signals—often reflecting noise or incidental patterns. Our goal is to preserve the former and perturb the latter, enabling contrastive views that are semantically aligned yet structurally diverse.

Frequency-based Graph Structure Decomposition We begin by applying a graph encoder to obtain node representations $\{\mathbf{h}_i\}_{i=1}^{|\mathcal{V}|}$, which capture local and global relational patterns for downstream spectral analysis . Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be the input graph. To quantify frequency contributions across different nodes and features, we introduce a frequency-sensitive smoothness measure based on pairwise feature variation:

$$S_p = 1 - \text{Normalize}\left(\frac{1}{|\mathcal{E}|}\sum_{(v_i,v_j)\in\mathcal{E}}\left(\mathbf{h}_i^{(p)} - \mathbf{h}_j^{(p)}\right)^2\right) \quad (1)$$

Here, $\mathbf{h}_i^{(p)}$ denotes the p-th feature of node v_i, and $S_p \in [0, 1]$ reflects the smoothness of that feature dimension, with higher values indicating lower frequency (i.e., global consistency). Using these scores, we define a node-level frequency contribution metric $K(v_i)$ as:

$$K(v_i) = \sum_{p=1}^{l}(1 - S_p)\sum_{(v_i,v_j)\in\mathcal{E}}\log\left(1 + \|\mathbf{h}_i^{(p)} - \mathbf{h}_j^{(p)}\|^2\right) \quad (2)$$

This formulation assigns larger values to nodes with more high-frequency behavior across less smooth features, capturing fine-grained structural fluctuations. The logarithmic scaling enhances stability while preserving discriminative signals. To partition the graph into frequency-aware substructures, we transform the continuous values $K(v_i)$ into binary hash codes via a nonlinear projection as follows:

$$B = \tanh(w \cdot K_v + b) \quad (3)$$

where w and b are learnable parameters. We then apply a symbolic function to classify the nodes. Specifically, if $\text{sign}(B_i) = -1$, node $v_i \in \mathcal{V}_L$ (low-frequency set); otherwise, $v_i \in \mathcal{V}_H$ (high-frequency set). An analogous procedure is used to divide the edge set into $\mathcal{E}_L$ and $\mathcal{E}_H$. To encourage meaningful spectral separation, we further regularize the decomposition by maximizing the divergence between low- and high-frequency components in the spectral embedding space:

$$\max \|\phi(\triangledown_{\text{low}}) - \phi(\triangledown_{\text{high}})\|_2^2 \quad (4)$$

where $\phi(\cdot)$ denotes a spectral embedding function , and $\triangledown_{\text{low}}$ and $\triangledown_{\text{high}}$ represent the subgraphs induced by low- and high-frequency nodes or edges, respectively.

Frequency-Driven Sample Generation

This section introduces the *Frequency-Driven Sample Generation* module, which utilizes the frequency-aware decomposition to guide augmentation in a principled manner. Based on the division of the graph into low- and high-frequency subgraphs, we design two complementary strategies: (i) for generating positive samples, we perturb only the high-frequency subgraph to inject localized variations while preserving the global structure captured by the low-frequency components; (ii) for negative sample generation, we gradually reducing the disturbance to the low-frequency components to construct hard negatives with increasing difficulty. This dual mechanism enhances both semantic consistency and discriminative power in graph contrastive learning.

Low-Frequency-Preserving Augmention The effectiveness of graph contrastive learning hinges on constructing augmented views that maintain semantic consistency while introducing sufficient diversity (Xia et al. 2025). To this end, we propose a *frequency-aware invariance* principle, where low-frequency components serve as proxies for semantic stability, and high-frequency structures are leveraged for controlled variability. Low-frequency signals typically correspond to the most stable and semantically relevant structures in a graph—they characterize global patterns and are naturally robust to perturbations. Based on this characteristic, we preserve the invariance of low-frequency components across enhanced views, thus retaining the semantically consistent parts of the graph. This design reflects the intuitive understanding that graph structure semantics often manifest as smooth low-frequency patterns. Therefore, by comparing the alignment of low-frequency subspaces between views, the model can learn semantically consistent and invariant representations, ultimately achieving frequency-guided semantic consistency. To implement this, we first derive binary masks for nodes and edges, $M_V \in \{0,1\}^{|\mathcal{V}|}$ and $M_E \in \{0,1\}^{|\mathcal{E}|}$, based on frequency contribution scores:

$$M_V[i] := \mathbb{1}_{[v_i \in \mathcal{V}_L]}, \quad M_E[u,v] := \mathbb{1}_{[(u,v)\in \mathcal{E}_L]}, \tag{5}$$

where $\mathbb{1}_{[\cdot]}$ is the indicator function selecting low-frequency components. Using these masks, We further decompose the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ into the following two parts :

$$\mathcal{G}_L = (\mathcal{V} \odot M_V, \mathcal{E} \odot M_E), \tag{6}$$

$$\mathcal{G}_H = (\mathcal{V} \odot (1 - M_V), \mathcal{E} \odot (1 - M_E)), \tag{7}$$

where $\odot$ denotes element-wise masking. Here, $\mathcal{G}_L$ encodes the semantic core of the graph, while $\mathcal{G}_H$ introduces informative high-frequency noise. For positive sample generation, we perturb only the high-frequency part:

$$\mathcal{G}^+ = \mathcal{G}_L \cup r(\mathcal{G}_H), \tag{8}$$

where $r(\mathcal{G}) \sim R(\mathcal{G})$ denotes a stochastic augmentation operator (e.g., node dropout, edge perturbation). This approach ensures that semantic integrity is preserved while encouraging meaningful view diversity.

Negative Sample Generation with Dynamic Difficulty Adjustment The discriminative strength of contrastive learning is closely tied to the difficulty of negative samples. Inspired by curriculum contrastive learning (Kalantidis et al. 2020), we present a dynamic difficulty adjustment strategy that gradually increases the hardness of negative samples over training. Specifically, we define negative sample $\mathcal{G}_i^-$ for each graph $\mathcal{G}_i \in \mathcal{B} = \{\mathcal{G}_1, \ldots, \mathcal{G}_N\}$ at epoch $t \in [0, T]$ as:

$$\mathcal{G}_i^- = \begin{cases} \{\mathcal{G}_j\}_{j\neq i} \subset \mathcal{B}, & \text{if } t \leq 0.5T, \\ H(\mathcal{G}_i; t), & \text{if } t > 0.5T, \end{cases} \tag{9}$$

where $H(\mathcal{G}_i; t)$ denotes a function that generates hard negatives by perturbing the low-frequency structure of $\mathcal{G}_i$. To ensure a smooth difficulty transition, we design a decaying removal schedule based on an exponential decay function:

$$Q_V(t) = \left[\alpha + (1-\alpha)\cdot e^{-10t/T}\right] \cdot |\mathcal{V}_L|, \tag{10}$$

$$Q_E(t) = \left[\alpha + (1-\alpha)\cdot e^{-10t/T}\right] \cdot |\mathcal{E}_L|, \tag{11}$$

where $\alpha \in (0, 1)$ is a hyperparameter controlling the perturbation ratio. Based on $Q_V(t)$ and $Q_E(t)$, we sample subsets of $\mathcal{V}_L$ and $\mathcal{E}_L$. Combine the sampling results with the high-frequency sets ($\mathcal{V}_H$ and $\mathcal{E}_H$) to generate the node mask M_V^- and the edge mask M_E^-. Use the generated masks to construct negative samples with increasing difficulty:

$$H(\mathcal{G}_i; t) = (\mathcal{V}_i \odot M_V^-, \mathcal{E}_i \odot M_E^-). \tag{12}$$

Through dynamic difficulty adjustment, our strategy encourages the model to learn increasingly discriminative features throughout training. By combining frequency-preserving positive augmentation and progressively harder negatives, this module significantly boosts the robustness and generalization of the learned graph representations.

Optimization Objective

The optimization objective plays a central role in graph contrastive learning, as it governs the training dynamics and determines the quality of learned representations. Our goal is to drive the model towards learning semantically meaningful and discriminative representations by maximizing the mutual information between different views of the same graph. Following this principle, we adopt the InfoNCE loss (Sun et al. 2020) as a tractable surrogate for mutual information maximization. InfoNCE encourages consistency between positive pairs while simultaneously enforcing discrimination from negative samples, thereby enhancing feature separability. Formally, let $\mathcal{B}$ denote a mini-batch of N input graphs. For each graph $\mathcal{G}_i \in \mathcal{B}$, we generate two stochastic positive view, $\mathcal{G}_i^1$ and $\mathcal{G}_i^2$, through low-frequency-preserving augmention strategies. At the same time, based on the current training epoch t, we use a dynamic difficulty adjustment strategy to dynamically generate negative samples $\mathcal{G}_i^-$. Both views are passed through a shared graph encoder $f(\cdot)$ to obtain their latent representations: $z_i^1 = f(\mathcal{G}_i^1)$, $z_i^2 = f(\mathcal{G}_i^2)$ and $z_i^- = f(\mathcal{G}_i^-)$. The contrastive objective is defined:

$$\mathcal{L}_{\text{Final}} = -\log \frac{\exp\big(\text{sim}(z_i^1, z_i^2)/\tau\big)}{\exp(\text{sim}(z_i^1, z_i^2)/\tau) + \exp\big(\text{sim}(z_i^1, z_i^-)/\tau\big)} \tag{13}$$

Here, $\text{sim}(\cdot,\cdot)$ denotes a similarity function, typically cosine similarity, and τ is a temperature scaling parameter that controls the sharpness of the softmax distribution. This objective promotes high similarity between positive pairs (z_i^1, z_i^2)—representations of the same graph under different augmentations—while discouraging similarity with negative pairs (z_i^1, z_i^-) . As a result, the model learns robust and discriminative representations that preserve the underlying graph semantics across augmentations.

Theoretical Analysis

Robustness Analysis

In practical deployment scenarios, GNNs are often exposed to various types of noise and perturbations. A robust model is expected to produce stable and consistent outputs in the presence of such disturbances. In this section, we provide a theoretical analysis of the robustness of the core formula of the proposed FA-GCL—i.e., Eq. (2). Specifically, we prove that Eq. (2) is c-Lipschitz continuous with respect to the features of the input nodes.This Lipschitz property ensures that the model's output does not vary excessively in response to small perturbations in the input, thereby enhancing its stability in noisy environments. To facilitate our analysis, we begin by presenting several fundamental definitions.

Definition 1 (**c-Lipschitz Continuity**). *A function $f : \mathcal{X} \to \mathcal{Y}$ is called c-*Lipschitz continuous *if there exists a constant $c > 0$ such that, for all $x_1, x_2 \in \mathcal{X}$,*

$$\|f(x_1) - f(x_2)\| \leq c \cdot \|x_1 - x_2\| .$$

Then the output of the function will not fluctuate excessively in response to small perturbations in the input.

Definition 2 (**Mean Value Theorem**). *For a differentiable function f on the interval $[a, b]$, there exists some $\xi \in (a, b)$ such that:*

$$f(b) - f(a) = f'(\xi) \cdot (b - a).$$

Definition 3 (**Triangle Inequality**). *For any vectors u, w, the following holds:*

$$\|u \pm w\| \leq \|u\| + \|w\| .$$

Based on these definitions, we now present the main robustness **Theorem 1** of FA-GCL.

Theorem 1. *Let $K(v_i)$ denote the frequency-aware contribution score of node v_i as defined in Eq. (2). For any perturbation Δh applied to the input features, there exists a constant:*

$$c = 2 \cdot \left(\max_{v_i \in \mathcal{V}} d_i \right) \cdot l, \tag{14}$$

such that:

$$\left| K\left(v_i; h + \Delta h\right) - K\left(v_i; h\right) \right| \leq c \cdot \|\Delta h\| , \tag{15}$$

where d_i denotes the degree of node v_i and l is the feature dimension. This result establishes that the Eq. (2) is theoretically robust to input noise.

Proof of Theorem 1

Next, we will verify the correctness of Theorem 1 through three steps. It should be noted that since all quantities involved in the derivation are real scalars, the Euclidean norm and absolute value are equivalent (i.e., $\|x\| = |x|$ for any $x \in \mathbb{R}$). Therefore, the use of either form does not affect the correctness of our analysis. This equivalence ensures consistency when verifying norm-based conditions.

Step 1: Lipschitz Continuity of the Inner Function Define the inner function in Eq. (2) as:

$$f\left(\varphi\right) = \log\left(1 + \varphi^2\right), \tag{16}$$

where $\varphi = \left\| h_i^{(p)} - h_j^{(p)} \right\|$ is the Euclidean distance between the p-th dimension features of nodes v_i and v_j. Next, we calculate the gradient of f :

$$f^{'}\left(\varphi\right) = \frac{2\varphi}{1 + \varphi^2}. \tag{17}$$

Since $f'(\varphi) \leq 1$ for all $\varphi \geq 0$ (achieved at $\varphi = 1$), by the Mean Value Theorem (Definition 2), for any $\varphi_1, \varphi_2 \geq 0$,

$$\begin{aligned} |f(\varphi_1) - f(\varphi_2)| &= |f'(\xi)| \cdot |\varphi_1 - \varphi_2| \\ &\leq |\varphi_1 - \varphi_2| , \end{aligned} \tag{18}$$

for some ξ between φ_1 and φ_2. Thus, we can conclude that f is 1-Lipschitz continuous.

Step 2: Noise Propagation in Feature Distances Assume that the graph data is noisy, resulting in changes to node features. Let $\Delta h_i^{(p)}$ and $\Delta h_j^{(p)}$ denote the noise in the p-th feature dimension of nodes v_i and v_j, respectively. The original feature distance φ changes to:

$$\varphi^{'} = \left\| \left(h_i^{(p)} + \Delta h_i^{(p)} \right) - \left(h_j^{(p)} + \Delta h_j^{(p)} \right) \right\| . \tag{19}$$

It follows that:

$$\begin{aligned} |\Delta\varphi| &= \left| \left\| \left(h_i^{(p)} + \Delta h_i^{(p)} \right) - \left(h_j^{(p)} + \Delta h_j^{(p)} \right) \right\| \right. \\ &\quad \left. - \left\| h_i^{(p)} - h_j^{(p)} \right\| \right| \\ &= \left| \left\| \left(h_i^{(p)} - h_j^{(p)} \right) + \left(\Delta h_i^{(p)} - \Delta h_j^{(p)} \right) \right\| \right. \\ &\quad \left. - \left\| h_i^{(p)} - h_j^{(p)} \right\| \right| . \end{aligned} \tag{20}$$

According to the triangle inequality, we have:

$$|\Delta\varphi| \leq \left\| \Delta h_i^{(p)} - \Delta h_j^{(p)} \right\| \leq \left\| \Delta h_i^{(p)} \right\| + \left\| \Delta h_j^{(p)} \right\| . \tag{21}$$

Combining this with the 1-Lipschitz continuity of the function $f(\varphi)$, the change in the inner function value satisfies:

$$\begin{aligned} |\Delta f| &= \left| f\left(\varphi'\right) - f\left(\varphi\right) \right| \\ &\leq |\Delta\varphi| \leq \left\| \Delta h_i^{(p)} \right\| + \left\| \Delta h_j^{(p)} \right\| . \end{aligned} \tag{22}$$

Step 3: Global Lipschitz Continuity of the Full Expression In Eq. (2),the frequency-aware contribution score $K(v_i)$ sums over neighbors and feature dimensions:

$$K(v_i) = \sum_{p=1}^{l} (1 - s_p) \sum_{(v_i, v_j) \in \mathcal{E}} f\left(\|h_i^{(p)} - h_j^{(p)}\| \right) . \tag{23}$$

As established in Step 2, when the features are perturbed by Δh, the variation in inner function f is bounded by:

$$|\Delta f| \leq \left\| \Delta h_i^{(p)} \right\| + \left\| \Delta h_j^{(p)} \right\| \leq 2\Delta h. \tag{24}$$

For each fixed p, the sum over d_i neighbors accumulates at most perturbation $2d_i\Delta h$. Since $(1 - s_p) \in [0, 1]$, it does not amplify the perturbation.Therefore, we can obtain:

$$\begin{aligned} \sum_{p=1}^{l} (1 - s_p) \sum_{(v_i, v_j) \in \mathcal{E}} |\Delta f| &\leq \sum_{p=1}^{l} 2 d_i \Delta h \\ &\leq 2 l d_i \Delta h. \end{aligned} \tag{25}$$

Taking the worst-case degree across all nodes, we obtain the global Lipschitz constant:

$$c = 2 \cdot \left(\max_{v_i \in \mathcal{V}} d_i \right) \cdot l. \tag{26}$$

Therefore, the total deviation in $K(v_i)$ due to input noise is bounded by:

$$\begin{aligned} |K(v_i; h+\Delta h) - K(v_i; h)| &= \left| \sum_{p=1}^{l} (1-s_p) \sum_{(v_i,v_j)\in\mathcal{E}} \Delta f \right| \\ &\le \sum_{p=1}^{l} (1-s_p) \sum_{(v_i,v_j)\in\mathcal{E}} |\Delta f| \\ &\le c \cdot \|\Delta h\| \end{aligned} \tag{27}$$

Experiments

In this section, we empirically evaluate the effectiveness of our proposed self-supervised framework, FA-GCL, in the context of graph representation learning. To assess its performance, we compare FA-GCL against representative and advanced methods under *unsupervised learning* settings for graph classification tasks. This experiments aim to highlight FA-GCL's ability to rapidly learn robust and informative representations from complex graph-structured data.

Category	Datasates	#Graphs	#Avg.Nodes	#Avg.Degree
Small molecules	MUTAG	188	17.93	19.79
Small molecules	NCI-1	4,110	29.87	32.30
Social	COLLAB	5,000	74.49	2457.78
Social	IMDB-B	1,000	19.77	96.53
Social	RDT-B	2,000	429.63	497.75
Social	RRDT-M5K	4,999	508.52	594.87
Bioinformatics	PROTEINS	1,113	39.06	72.82
Bioinformatics	D&D	1,178	284.32	715.66

Table 1: Statistics for unsupervised learning datasets.

Datasets We evaluate our method on total eight widely used real-world datasets[1], spanning three application domains, as shown in **Table 1**. Specifically, we use two molecular datasets (MUTAG, NCI1), four social network datasets (COLLAB, RDT-B, RDT-M5K, IMDB-B), and two bioinformatics datasets (PROTEINS, D&D).

Baseline Methods We primarily compare FA-GCL against representative and advanced graph contrastive learning methods, including: InfoGraph (Sun et al. 2020), GraphCL (You et al. 2020), JOAO (You et al. 2021), AD-GCL (Suresh et al. 2021), SimGRACE (Xia et al. 2022), LG2AR (Hassani and Ahmadi 2022), RGCL (Li et al. 2022), GCS (Wei et al. 2023), DRGCL (Ji et al. 2024), DGPM (Yan et al. 2024), CI-GCL (Tan et al. 2024),and Dual-Prism (Xia et al. 2025).

Parameter Settings We follow standard evaluation protocols adopted by prior works (Chu et al. 2021; You et al. 2020; Sun et al. 2020; Narayanan et al. 2017) to ensure fair and consistent comparisons. All models are evaluated using 10-fold cross-validation, and we report the average classification accuracy across 10 runs as the primary metric. For supervised baselines, we adopt the reported results from original papers when applicable; otherwise, we re-implement and tune them based on their released configurations. Classification is performed using LIBSVM (Chang and Lin 2011), where the regularization parameter C is selected from the set $\{10^{-3}, 10^{-2}, \ldots, 10^{2}, 10^{3}\}$. For FA-GCL, we use a 4-layer GIN (Xu et al. 2019) as the backbone encoder to maintain consistency across baselines. The hidden dimension is set to 128, batch size to 128, learning rate to 0.001, and the final graph-level embedding dimension to 32.

[1]The datasets are publicly available at https://ls11-www.cs.tu-dortmund.de/staff/morris/graphkerneldatasets.

Overall Comparison

Table 2 presents the classification results on multiple benchmark datasets. Several key observations emerge:

- **Superior and consistent performance.** FA-GCL consistently achieves state-of-the-art results across all datasets. Even on strong-performing datasets such as MUTAG and NCI1, where existing methods already yield high accuracy, FA-GCL further improves performance. This highlights the robustness of our frequency-aware design in capturing structural nuances that are critical for graph classification.
- **Clear advantage over recent baselines.** Compared to strong unsupervised approaches such as GCS and DGPM, FA-GCL achieves higher accuracy across diverse graph domains. While these baselines rely on sophisticated rule-based augmentations, our frequency-aware perturbation adaptively targets high-frequency components, enabling the extraction of informative local features while preserving global semantics. This advantage is particularly evident on structurally complex datasets such as D&D and RDT-B.
- **Effectiveness of frequency-guided perturbation.** FA-GCL's performance gains stem from its ability to precisely identify and perturb frequency-aware structures. Unlike GCS, which uses semantic node sampling, FA-GCL directly leverages spectral decomposition to operate on high-frequency signals, capturing fine-grained features that enhance contrastive learning effectiveness.
- **Improvement over frequency-domain methods.** FA-GCL surpasses representative spectral approaches such as CI-GCL and Dual-Prism. This improvement is attributed to our learnable frequency-aware decomposition module, which adaptively quantifies node- and edge-level frequency contributions and dynamically modulates perturbation intensity. This design generates more informative and generalizable representations than static spectral methods.

Ablation Study

We conduct ablation studies on three representative datasets (MUTAG, PROTEINS, and IMDB-B) to evaluate the contributions of two key components: *low-frequency-preserving augmentation* and *dynamic difficulty adjustment*. Specifically, we compare the FA-GCL variants:

Model	Small Molecules		Social Network				Bioinformatics	
	MUTAG	NCI1	COLLAB	IMDB-B	RDT-B	RDT-M5K	PROTEINS	D&D
InfoGraph (NeurIPS'20)	89.0±1.1	76.2±1.1	70.7±1.1	73.0±0.9	82.5±1.4	53.5±1.0	74.4±0.3	72.9±1.8
GraphCL (NeurIPS'20)	86.8±1.3	77.9±0.4	71.4±1.2	71.1±0.4	89.5±0.8	56.0±0.3	74.4±0.5	78.6±0.4
JOAO (ICML'21)	87.7±0.8	73.0±0.8	70.4±2.2	71.6±0.9	78.4±1.4	45.6±2.9	71.3±0.9	66.9±1.8
AD-GCL (NeurIPS'21)	88.6±1.2	75.8±0.6	73.9±0.7	71.5±1.0	85.3±1.0	54.7±0.6	75.0±0.5	75.8±0.4
SimGRACE (WWW'22)	89.0±1.3	79.1±0.4	72.4±1.2	71.3±0.7	86.2±1.1	50.6±1.9	75.3±0.1	76.2±1.3
LG2AR (ICML'22)	90.0±0.6	75.6±0.9	72.8±0.7	74.5±0.6	91.8±0.4	53.7±0.8	75.0±0.5	79.1±0.3
RGCL (ICML'22)	87.7±1.0	78.1±1.1	70.9±1.8	71.9±0.8	90.3±0.6	56.4±0.4	75.2±0.4	78.9±0.5
GCS (ICML'23)	88.1±0.7	76.2±0.9	74.2±0.7	72.3±0.6	90.2±0.8	56.3±0.3	75.1±0.5	76.4±0.3
DRGCL(AAAI'24)	89.5±0.6	78.7±0.4	70.6±1.4	72.0±0.5	90.8±0.3	56.3±0.2	75.2±0.6	78.4±0.7
DGPM (AAAI'24)	90.4±0.8	78.5±0.7	75.6±0.9	74.6±0.3	87.7±0.5	56.0±0.4	74.2±0.5	77.6±0.2
CI-GCL (ICML'24)	89.6±0.9	80.5±0.5	70.7±0.7	73.8±0.8	90.8±0.5	56.5±0.3	76.5±0.1	79.6±0.3
Dual-Prism (AAAI'25)	89.9±1.3	79.4±0.2	73.2±0.7	71.7±0.3	91.2±0.2	55.9±0.4	74.7±0.2	79.9±1.0
Ours	**94.5±1.1**	**81.2±0.6**	**77.9±0.5**	**77.2±0.7**	**93.3±0.6**	**57.6±1.2**	**79.6±0.8**	**81.5±1.6**

Table 2: Unsupervised representation learning classification accuracy (%) on TU datasets. **Bold** denotes the best performance while Underline represents the second best performance.

Variation	MUTAG	PROTEINS	IMDB-B
FA-GCL	94.5%	79.6%	77.2%
w/o low-frequency-preserving aug	89.8%	75.4%	74.1%
w/o dynamic difficulty adjustment	90.7%	76.1%	73.7%

Table 3: Ablation study on the key components of our method. The term "w/o" indicates "without".

- **w/o low-frequency-preserving augmentation**: replaces the frequency-aware augmentation with random edge/node deletion to assess the effect of high-frequency perturbation.
- **w/o dynamic difficulty adjustment**: substitutes the adaptive negative sampling with random negatives, to evaluate the effect of hardness-aware contrastive learning.

As reported in **Table 3**, both components significantly influence model performance. On MUTAG, removing low-frequency-preserving augmentation causes accuracy drop, demonstrating its critical role in generating informative positive views. Similarly, excluding dynamic difficulty adjustment results in a accuracy decrease, underscoring the importance of adaptive hard negative sampling. Consistent trends are observed on PROTEINS and IMDB-B, where the full FA-GCL model outperforms each variant. These results validate our design, indicating that frequency-aware augmentation and hardness-aware contrastive sampling are individually effective and mutually complementary. The low-frequency-preserving augmentation maintains semantic consistency across views, enabling the model to focus on globally relevant structures, while dynamic difficulty negatives promote more discriminative and fine-grained representations.

Hyperparameter Analysis

To gain deeper insights into the internal mechanism of FA-GCL, we conduct a systematic sensitivity analysis on a key hyperparameter α, which controls the minimum proportion of low-frequency structural components to be removed during the generation of hard negative samples. This parameter directly influences the overall semantic difficulty of the negative samples and thus plays a critical role in contrastive learning. We evaluate the impact of α on three representative graph datasets: MUTAG, PROTEINS, and IMDB-B. The hyperparameter is varied in the range $\alpha \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, with all other configurations held constant. The results are summarized in **Fig. 3**.

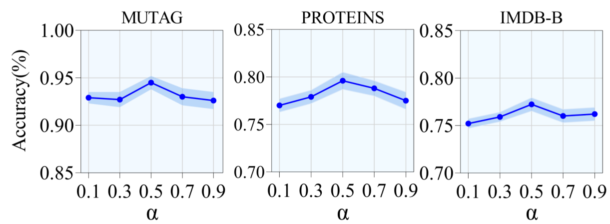

Figure 3: Hyperparameter Analysis

The experimental results demonstrate that the model achieves optimal performance when $\alpha = 0.5$. A too-small α retains excessive low-frequency structure in the late training stages, leading to negative samples that are semantically too close to the original graphs, thereby weakening the contrastive signal. On the other hand, an excessively large α severely limits the adjustable range for the difficulty of negative samples, resulting in insufficient progression in difficulty throughout training and ultimately restricting the model's capacity to learn complex structural features. Overall, a moderate setting of α strikes a favorable balance between preserving semantic stability and introducing structural perturbation. This enables more effective contrastive learning by gradually increasing negative sample difficulty, thereby maximizing the model's generalization performance.

Conclusion

In this study, we propose **FA-GCL**, an adaptive frequency-aware GCL framework that identifies and modulates high- and low-frequency structural components in graphs. By preserving low-frequency semantics and perturbing high-frequency signals, FA-GCL generates more informative positives and progressively harder negatives to improve representation learning. As frequency-aware GCL remains underexplored, our work highlights its potential and opens a promising direction for building robust, and semantically graph representations in real-world scenarios.

Acknowledgments

This work was supported by Qingdao Municipal Natural Science Foundation (Grant No. 25-1-1-107-zyyd-jch), the National Natural Science Foundation of China (Grant No.62502113), Shandong Provincial Natural Science Foundation of China (Grant No. ZR2025QC640), Hainan Provincial Natural Science Foundation of China (Grant No. 625QN383), Heilongjiang Provincial Natural Science Foundation of China (Grant No.QC2025F011). Additionally, this work was also supported by Fund of Intelligent Internet of Things and Crowd Computing (No. B25029), Natural Science Foundation of Shandong Province (No. ZR2024QF177), Hunan Provincial Natural Science Foundation (No. 2025JJ60393), Changsha Natural Science Foundation (No. 106271), Xiangjiang Laboratory General Project (No. 25XJ03016, 23XJ01001, 22XJ01001, 24XJJCYJ01002), the Fundamental Research Funds for the Central Universities (No. GK762026011560, GK762026011561) at Harbin Engineering University.

References

Cao, D.; Wang, Y.; Duan, J.; Zhang, C.; Zhu, X.; Huang, C.; Tong, Y.; Xu, B.; Bai, J.; Tong, J.; and Zhang, Q. 2021. Spectral Temporal Graph Neural Network for Multivariate Time-series Forecasting. *CoRR*, abs/2103.07719.

Chang, C.; and Lin, C. 2011. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3): 27:1–27:27.

Chang, H.; Rong, Y.; Xu, T.; Bian, Y.; Zhou, S.; Wang, X.; Huang, J.; and Zhu, W. 2021. Not All Low-Pass Filters are Robust in Graph Convolutional Networks. In Ranzato, M.; Beygelzimer, A.; Dauphin, Y. N.; Liang, P.; and Vaughan, J. W., eds., *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, 25058–25071.

Chu, G.; Wang, X.; Shi, C.; and Jiang, X. 2021. CuCo: Graph Representation with Curriculum Contrastive Learning. In Zhou, Z., ed., *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI 2021, Virtual Event / Montreal, Canada, 19-27 August 2021*, 2300–2306. ijcai.org.

Hassani, K.; and Ahmadi, A. H. K. 2022. Learning Graph Augmentations to Learn Graph Representations. *CoRR*, abs/2201.09830.

Ji, Q.; Li, J.; Hu, J.; Wang, R.; Zheng, C.; and Xu, F. 2024. Rethinking Dimensional Rationale in Graph Contrastive Learning from Causal Perspective. In Wooldridge, M. J.; Dy, J. G.; and Natarajan, S., eds., *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2014, February 20-27, 2024, Vancouver, Canada*, 12810–12820. AAAI Press.

Jiang, P.; Xiao, C.; Fu, T.; Bhatia, P.; Kass-Hout, T. A.; Sun, J.; and Han, J. 2025. Bi-level Contrastive Learning for Knowledge-Enhanced Molecule Representations. In Walsh, T.; Shah, J.; and Kolter, Z., eds., *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence, February 25 - March 4, 2025, Philadelphia, PA, USA*, 352–360. AAAI Press.

Kalantidis, Y.; Sariyildiz, M. B.; Pion, N.; Weinzaepfel, P.; and Larlus, D. 2020. Hard Negative Mixing for Contrastive Learning. In Larochelle, H.; Ranzato, M.; Hadsell, R.; Balcan, M.; and Lin, H., eds., *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*.

Li, S.; Wang, X.; Zhang, A.; Wu, Y.; He, X.; and Chua, T. 2022. Let Invariant Rationale Discovery Inspire Graph Contrastive Learning. In Chaudhuri, K.; Jegelka, S.; Song, L.; Szepesvári, C.; Niu, G.; and Sabato, S., eds., *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, 13052–13065. PMLR.

Lin, L.; Chen, J.; and Wang, H. 2023. Spectral Augmentation for Self-Supervised Learning on Graphs. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net.

Narayanan, A.; Chandramohan, M.; Venkatesan, R.; Chen, L.; Liu, Y.; and Jaiswal, S. 2017. graph2vec: Learning Distributed Representations of Graphs. *CoRR*, abs/1707.05005.

Peng, W.; Zhou, Z.; Dai, W.; Yu, N.; and Wang, J. 2024. Multi-Network Graph Contrastive Learning for Cancer Driver Gene Identification. *IEEE Trans. Netw. Sci. Eng.*, 11(4): 3430–3440.

Qin, Y.; Ju, W.; Luo, X.; Gu, Y.; Xiao, Z.; and Zhang, M. 2024. PolyCF: Towards the Optimal Spectral Graph Filters for Collaborative Filtering. *CoRR*, abs/2401.12590.

Shen, Y.; Jiang, X.; Li, Z.; Wang, Y.; Xu, C.; Shen, H.; and Cheng, X. 2023. UniSKGRep: A unified representation learning framework of social network and knowledge graph. *Neural Networks*, 158: 142–153.

Sun, F.; Hoffmann, J.; Verma, V.; and Tang, J. 2020. InfoGraph: Unsupervised and Semi-supervised Graph-Level Representation Learning via Mutual Information Maximization. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net.

Suresh, S.; Li, P.; Hao, C.; and Neville, J. 2021. Adversarial Graph Augmentation to Improve Graph Contrastive Learning. In Ranzato, M.; Beygelzimer, A.; Dauphin, Y. N.; Liang, P.; and Vaughan, J. W., eds., *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, 15920–15933.

Tan, S.; Li, D.; Jiang, R.; Zhang, Y.; and Okumura, M. 2024. Community-Invariant Graph Contrastive Learning. In *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*. OpenReview.net.

Wei, C.; Wang, Y.; Bai, B.; Ni, K.; Brady, D.; and Fang, L. 2023. Boosting Graph Contrastive Learning via Graph Contrastive Saliency. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *International*

Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA, volume 202 of *Proceedings of Machine Learning Research*, 36839–36855. PMLR.

Xia, J.; Wu, L.; Chen, J.; Hu, B.; and Li, S. Z. 2022. SimGRACE: A Simple Framework for Graph Contrastive Learning without Data Augmentation. In Laforest, F.; Troncy, R.; Simperl, E.; Agarwal, D.; Gionis, A.; Herman, I.; and Médini, L., eds., *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, 1070–1079. ACM.

Xia, Y.; Yu, R.; Liang, Y.; Bresson, X.; Wang, X.; and Zimmermann, R. 2025. Through the Dual-Prism: A Spectral Perspective on Graph Data Augmentation for Graph Classifications. In Walsh, T.; Shah, J.; and Kolter, Z., eds., *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence, February 25 - March 4, 2025, Philadelphia, PA, USA*, 21635–21643. AAAI Press.

Xiao, Z.; Wang, H.; Jin, Y.; Feng, L.; Chen, G.; Huang, F.; and Zhao, J. 2023. SPA: A Graph Spectral Alignment Perspective for Domain Adaptation. In Oh, A.; Naumann, T.; Globerson, A.; Saenko, K.; Hardt, M.; and Levine, S., eds., *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*.

Xu, K.; Hu, W.; Leskovec, J.; and Jegelka, S. 2019. How Powerful are Graph Neural Networks? In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net.

Xu, Y.; Huang, S.; Zhang, H.; and Li, X. 2025. Why Does Dropping Edges Usually Outperform Adding Edges in Graph Contrastive Learning? In Walsh, T.; Shah, J.; and Kolter, Z., eds., *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence, February 25 - March 4, 2025, Philadelphia, PA, USA*, 21824–21832. AAAI Press.

Yan, P.; Song, K.; Jiang, Z.; Kang, Y.; Lin, T.; Sun, C.; and Liu, X. 2024. Empowering Dual-Level Graph Self-Supervised Pretraining with Motif Discovery. In Wooldridge, M. J.; Dy, J. G.; and Natarajan, S., eds., *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2014, February 20-27, 2024, Vancouver, Canada*, 9223–9231. AAAI Press.

You, Y.; Chen, T.; Shen, Y.; and Wang, Z. 2021. Graph Contrastive Learning Automated. In Meila, M.; and Zhang, T., eds., *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, 12121–12132. PMLR.

You, Y.; Chen, T.; Sui, Y.; Chen, T.; Wang, Z.; and Shen, Y. 2020. Graph Contrastive Learning with Augmentations. In Larochelle, H.; Ranzato, M.; Hadsell, R.; Balcan, M.; and Lin, H., eds., *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*.

Zhang, H.; Ren, Y.; Fu, L.; Wang, X.; Chen, G.; and Zhou, C. 2024. Multi-Scale Self-Supervised Graph Contrastive Learning With Injective Node Augmentation. *IEEE Trans. Knowl. Data Eng.*, 36(1): 261–274.

Zheng, X.; Feng, X.; Wu, B.; Ren, B.; Li, W.; Hao, X.; and Liang, X. 2025. Breaking Semantic Barriers: A Zero-Shot Generalized Framework for Graph Anomaly Detection. In *Proceedings of the 33nd ACM International Conference on Multimedia, MM 2025, 27 October 2025 - 1 November 2025*, 1423–1432. ACM.

Zheng, X.; Hao, X.; Wu, B.; Bao, X.; Zhang, X.; Li, W.; and Liang, X. 2024a. A Sample-driven Selection Framework: Towards Graph Contrastive Networks with Reinforcement Learning. In Cai, J.; Kankanhalli, M. S.; Prabhakaran, B.; Boll, S.; Subramanian, R.; Zheng, L.; Singh, V. K.; César, P.; Xie, L.; and Xu, D., eds., *Proceedings of the 32nd ACM International Conference on Multimedia, MM 2024, Melbourne, VIC, Australia, 28 October 2024 - 1 November 2024*, 10755–10764. ACM.

Zheng, X.; Liang, X.; Wu, B.; Guo, Y.; and Zhang, X. 2022. Graph Capsule Network with a Dual Adaptive Mechanism. In Amigó, E.; Castells, P.; Gonzalo, J.; Carterette, B.; Culpepper, J. S.; and Kazai, G., eds., *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, 1859–1864. ACM.

Zheng, X.; Wu, B.; Liang, X.; and Li, W. 2024b. Friend or Foe? Mining Suspicious Behavior via Graph Capsule Infomax Detector against Fraudsters. In Chua, T.; Ngo, C.; Kumar, R.; Lauw, H. W.; and Lee, R. K., eds., *Proceedings of the ACM on Web Conference 2024, WWW 2024, Singapore, May 13-17, 2024*, 2684–2693. ACM.

Zhu, Y.; Xu, Y.; Yu, F.; Liu, Q.; Wu, S.; and Wang, L. 2020. Deep Graph Contrastive Representation Learning. *CoRR*, abs/2006.04131.