

# HTG-GCL: Leveraging Hierarchical Topological Granularity from Cellular Complexes for Graph Contrastive Learning

Qirui Ji<sup>1 2 \*</sup>, Bin Qin<sup>1 2 \*</sup>, Yifan Jin<sup>1 2</sup>, Yunze Zhao<sup>1 2</sup>,  
Chuxiong Sun<sup>1</sup>, Changwen Zheng<sup>1</sup>, Jianwen Cao<sup>1</sup>, Jiangmeng Li<sup>1 †</sup>

<sup>1</sup>National Key Laboratory of Space Integrated Information System, Institute of Software Chinese Academy of Sciences

<sup>2</sup>University of Chinese Academy of Sciences

{jqirui2022, yifan2020, chuxiong2016, changwen, jiangmeng2019, jianwen}@iscas.ac.cn,  
{qinbin21, zhaoyunze19}@mails.ucas.ac.cn

## Abstract

Graph contrastive learning (GCL) aims to learn discriminative semantic invariance by contrasting different views of the same graph that share critical topological patterns. However, existing GCL approaches with structural augmentations often struggle to identify task-relevant topological structures, let alone adapt to the varying coarse-to-fine topological granularities required across different downstream tasks. To remedy this issue, we introduce *Hierarchical Topological Granularity Graph Contrastive Learning* (HTG-GCL), a novel framework that leverages transformations of the same graph to generate multi-scale ring-based cellular complexes, embodying the concept of topological granularity, thereby generating diverse topological views. Recognizing that a certain granularity may contain misleading semantics, we propose a multi-granularity decoupled contrast and apply a granularity-specific weighting mechanism based on uncertainty estimation. Comprehensive experiments on various benchmarks demonstrate the effectiveness of HTG-GCL, highlighting its superior performance in capturing meaningful graph representations through hierarchical topological information.

**Code** — <https://github.com/ByronJi/HTG-GCL>

## 1 Introduction

Graph Contrastive Learning (GCL) has emerged as an effective paradigm for learning discriminative graph representations by contrasting multiple views of the same graph, offering significant promise in domains such as biochemical molecular classification and social network analysis (Ju et al. 2024). Recent advances in structure-aware GCL methods (You et al. 2020; Suresh et al. 2021; You et al. 2021; Gao et al. 2022; Li et al. 2022, 2024) demonstrate the power of leveraging structural perturbations to generate informative views, enabling the model to effectively learn from unlabeled and non-Euclidean data. However, as complex data from natural systems becomes increasingly prevalent, representing such data with simple pairwise relationships in

graphs often falls short in capturing their inherent relational complexity. To model these higher-order relational structures, more expressive mathematical objects beyond graphs, such as simplicial complexes and cellular complexes, have been introduced as generalized data representations (Madhu and Chepuri 2023; Qin et al. 2025). These topological abstractions allow for a richer encoding of multi-way interactions and geometric constraints, thus providing a principled way to incorporate higher-order topology into representation learning frameworks.

However, current state-of-the-art GCL work (Qin et al. 2025) adopts a fixed ring size when constructing its cellular complexes, thereby capturing only a single level of topological *granularity*. Here, the *granularity* of a cellular complex refers to the scale at which the graph’s topological structure is represented, determined by the maximum ring size  $m$  used to attach 2-cells (see Section 4.1 for details). Such a static design assumes that a single topological resolution is universally effective across diverse domains—a premise that may not hold in practice. Different real-world scenarios often depend on distinct types of topological patterns. This observation is supported by our intuitive motivation experiments conducted in a supervised setting on the molecular dataset NCI1 and the social network dataset IMDB-B (Figure 1), which show that indiscriminately incorporating high-order structures may degrade performance when the extracted patterns do not align with the structural semantics of a specific domain. For instance, social network analysis typically benefits from identifying triangle structures (i.e., 3-cliques), and thus operates effectively at a coarse topological granularity that captures small rings. In contrast, molecular datasets, such as those used for toxicity prediction or mutagenicity detection, often require identifying larger and more complex ring structures, corresponding to a finer topological granularity. This divergence reflects the need for scenario-dependent topological granularity.

Yet in the context of GCL, the model lacks access to task-specific labels, making it difficult to determine which topological granularity is most relevant for downstream performance. A natural solution is to incorporate multiple levels of granularity—ranging from coarse to fine—so that the model can leverage a richer hierarchy of topological signals. This

\*These authors contributed equally.

†Corresponding author.

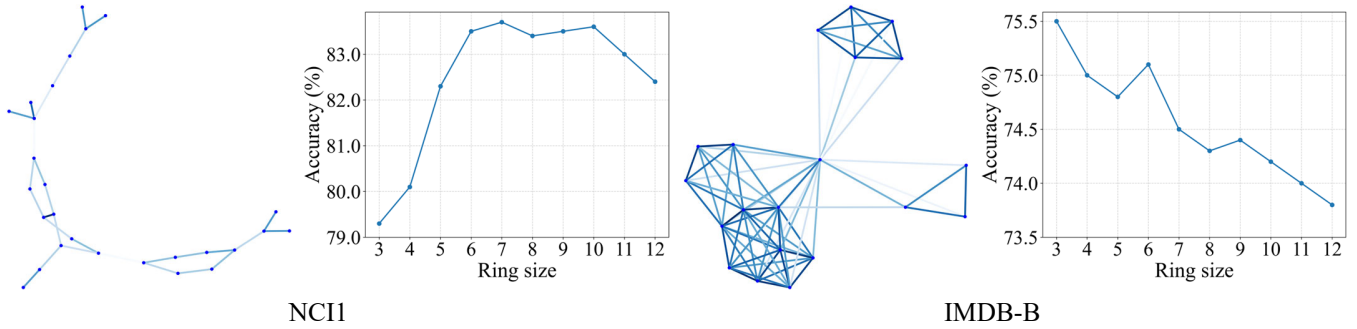


Figure 1: Visualization of graph structures (Zhao, Ying, and Yu 2024) and classification performance across different ring sizes on two TU datasets: NCI1 and IMDB-B. The results reveal that topological granularity significantly affects downstream performance, and indiscriminate inclusion of high-order structures may harm accuracy when they misalign with domain-specific semantics. For example, IMDB-B benefits from smaller rings (triangles), while NCI1 favors larger ring structures, highlighting the need for scenario-adaptive granularity modeling.

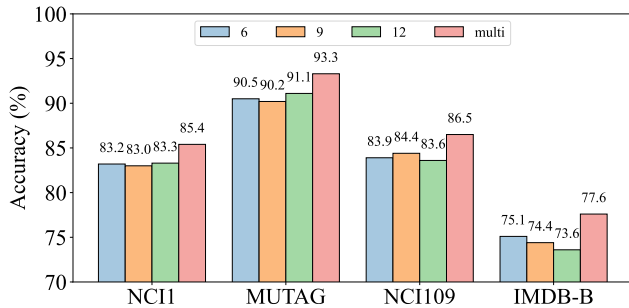


Figure 2: Performance comparison on four TU datasets under different single-ring granularities (6, 9, 12) and their multi-granularity integration. The multi-granularity setting consistently outperforms any single fixed granularity, demonstrating the complementary nature of hierarchical topological views.

motivates the following central question:

**“Can hierarchical multi-granularity topological modeling enhance downstream performance in GCL?”**

With the above question in mind, we conduct an additional motivation experiment on four TU datasets (NCI1, MUTAG, NCI109, and IMDB-B) to explore the impact of multi-granularity topological modeling on downstream classification tasks. As illustrated in Figure 2, we compare the classification accuracy achieved using cellular complexes constructed with different maximum ring sizes (6, 9, and 12), as well as a multi-granularity integration strategy that combines all three. The results consistently show that the multi-granularity approach outperforms any single fixed granularity across all datasets, indicating that integrating hierarchical topological information provides complementary structural cues that enhance representation learning.

These findings motivate us to develop a principled approach that can effectively leverage hierarchical topological information in a self-supervised setting. To this end, we introduce *Hierarchical Topological Granularity Graph Contrastive Learning* (HTG-GCL), a novel framework that

constructs diverse topological views by transforming the same graph into a series of ring-based cellular complexes with varying maximum ring sizes, thereby capturing multiple levels of topological granularity. Motivated by the observation that different scenarios may favor different granularities, we recognize that not all topological views are equally informative for every downstream task. To address this, we propose a *multi-granularity decoupled contrastive learning strategy* (MGDC), where each view is projected into a granularity-specific embedding space, under the hypothesis that certain particular granularity may be uninformative or even misleading. Moreover, since this hypothesis may not universally hold, we further incorporate an *uncertainty-based weighting mechanism* to assess the reliability of each granularity-specific space and adaptively reweight their contributions during contrastive training. Our contributions are summarized as follows:

- We introduce the *first* work that leverages cellular complexes to construct multi-granularity topological views for GCL, effectively capturing hierarchical topological information at different ring scales, which is a complementary perspective to traditional augmentations.
- We propose HTG-GCL, which generates multi-granularity topological views and incorporates MGDC, an uncertainty-weighted contrastive strategy over decoupled granularity-specific embedding spaces.
- We empirically validate the effectiveness of HTG-GCL through extensive experiments on multiple benchmarks.

## 2 Related Work

### 2.1 Graph Contrastive Learning

Recent advancements in graph contrastive learning have primarily focused on designing effective structural augmentation strategies and contrastive objectives to enhance graph-level representations. GraphCL (You et al. 2020) initiates this line of research by proposing four general augmentations to enforce consistency between perturbed graph views. To improve augmentation quality, ADGCL (Suresh et al. 2021) introduces adversarial strategies that avoid redundant

semantics, while JOAO (You et al. 2021) tackles the problem of augmentation selection via a min-max bi-level optimization framework. RGCL (Li et al. 2022) incorporates rationale discovery to preserve meaningful structural patterns, and SimGRACE (Xia et al. 2022) avoids handcrafted augmentations altogether by perturbing the encoder with Gaussian noise. HTML (Li et al. 2024) leverages knowledge distillation to align graph-level and subgraph-level views based on topological isomorphism. TopoSRL (Madhu and Chepuri 2023) introduces self-supervised learning on simplicial complexes with topology-preserving augmentations to capture higher-order interactions. CellCLAT (Qin et al. 2025) goes beyond standard graphs by introducing cellular complexes, aiming to enrich higher-order structural information while addressing topological redundancy. However, current methods lack the capacity to capture topological semantics at varying granularities. HTG-GCL addresses this by constructing multi-scale cellular complexes and utilizing granularity-decoupled contrastive learning.

## 2.2 Higher-order Representations of Graphs

Advancements in topological deep learning have shifted the focus from traditional graphs to more expressive structures, enabling the modeling of higher-order relationships (Papillon et al. 2024). Simplicial complexes capture hierarchical connectivity by connecting nodes with edges, faces, and higher-dimensional simplices. Models such as SIN (Bodnar et al. 2021b) and SCNN (Yang, Isufi, and Leus 2022) are well-suited for encoding these structures. Cellular complexes further generalize simplicial complexes by allowing arbitrary shapes for faces and volumes, making them suitable for representing more intricate topological structures. Cellular Complex Neural Networks (Hajij, Istvan, and Zamzmi 2020; Bodnar et al. 2021a) have been proposed to effectively encode such data.

## 3 Background

### 3.1 Cellular Complex

A  $d$ -dimensional cellular complex  $X$  is a topological space constructed inductively by a sequence of increasing  $k$ -skeleton:  $X^{(-1)} = \emptyset \subset X^{(0)} \subset X^{(1)} \subset \dots \subset X^{(d)} = X$ . The construction (Hansen and Ghrist 2019) begins with a discrete set of points, called 0-cells, which form the 0-skeleton  $X^{(0)}$ . Then, for  $k \geq 1$ , the  $k$ -skeleton  $X^{(k)}$  is formed by attaching  $k$ -cells  $\sigma$  (topological spaces homeomorphic to the open  $k$ -dimensional ball  $B^k = \{x \in \mathbb{R}^k : \|x\| < 1\}$ ) to the  $(k-1)$ -skeleton  $X^{(k-1)}$ . This attachment is guided by continuous map called **attaching map**:  $\varphi_\sigma^k: \mathbb{S}^{k-1} \rightarrow X^{(k-1)}$ , where  $(k-1)$ -sphere  $\mathbb{S}^{k-1}$  is the **boundary** ( $\partial\sigma \cong \mathbb{S}^{k-1}$ ) of the  $k$ -cell  $\sigma$  and the index  $\sigma$  ranges over all  $k$ -cells being attached. Finally, the entire cellular complex  $X$  is the union of all its skeletons  $X = \bigcup_{k=0}^d X^{(k)}$ .

**Example 1 (Graph as a 1-dim cellular complex):** Graph  $G = (V, E)$  is a 1-dim cellular complex  $G = X^{(1)}$ , where vertices  $V$  are 0-cells that constitute the 0-skeleton  $X^{(0)}$ .

Each edge  $e = \{u, v\} \in E$  is regarded as a 1-cell  $\sigma_e$  by attaching the endpoints of line segments to these vertices  $u, v$ . Thus, the 1-skeleton  $X^{(1)}$  is precisely the graph  $G$ .

**Example 2 (2-dim cellular complex from cycles):** Starting from a graph  $G$ , we attach a 2-cell  $\sigma_C$  (homeomorphic to a 2-dim disk) to an induced cycle  $C = (v_1, v_2, \dots, v_n, v_1) \subseteq G$  along its boundary circle with the cyclic sequence of edges forming  $C$ . The resulting 2-skeleton  $X^{(2)}$  is thus formed by filling in induced cycles of the original graph.

### 3.2 Cellular Adjacencies

We naturally extend the adjacency of nodes (0-cells) on graphs to hierarchical cross-dimensional higher-order interactions between edges (1-cells) and faces (2-cells). First, analogous to the incidence matrix on graphs, the incidence matrix between  $(k-1)$ -cells and  $k$ -cells is  $\mathbf{B}_k \in \mathbb{R}^{N_{k-1} \times N_k}$ , where  $N_k$  denotes the number of  $k$ -cells in  $X$ . The  $(i, j)$ -th entry of  $\mathbf{B}_k$  is 1 if the  $i$ -th  $(k-1)$ -cell is on the boundary of the  $j$ -th  $k$ -cell.

Based on the incidence matrix, we define four types of cellular adjacencies: 1) Boundary adjacent:  $\mathbf{A}_k^B = \mathbf{B}_{k+1}^T \in \mathbb{R}^{N_{k+1} \times N_k}$ ; 2) Co-boundary adjacent:  $\mathbf{A}_k^C = \mathbf{B}_k \in \mathbb{R}^{N_{k-1} \times N_k}$ ; 3) Upper adjacent:  $\mathbf{A}_k^U = \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T \in \mathbb{R}^{N_k \times N_k}$ ; 4) Lower adjacent:  $\mathbf{A}_k^L = \mathbf{B}_k^T \mathbf{B}_k \in \mathbb{R}^{N_k \times N_k}$ . The  $k$ -th Hodge Laplacian is  $\mathbf{L}_k = \mathbf{A}_k^U + \mathbf{A}_k^L$ . In particular, when  $k = 0$ ,  $\mathbf{L}_0 = \mathbf{B}_1 \mathbf{B}_1^T = \mathbf{D} - \mathbf{A}$  is the usual graph Laplacian. The adjacency matrix of a graph is a special case of cellular adjacencies, namely the upper adjacent  $\mathbf{A}_0^U$  of nodes (0-cells). Finally, we use the notation  $\mathcal{B}(\sigma)$ ,  $\mathcal{C}(\sigma)$ ,  $\mathcal{N}_U(\sigma)$ , and  $\mathcal{N}_L(\sigma)$  to denote the boundary, co-boundary, upper and lower neighbors of cell  $\sigma$ , respectively.

### 3.3 Cellular Complex Neural Networks

Let the initial features of the  $k$ -cells in a cellular complex  $X$  be denoted as  $\mathbf{X}_k \in \mathbb{R}^{N_k \times F_k}$ . Cellular Complex Neural Networks (CCNNs) encode these  $k$ -cells and learn their representations by aggregating messages from four types of neighbors. Specifically, the message-passing paradigm (Hajij, Istvan, and Zamzmi 2020; Bodnar et al. 2021a) for learning the  $l$ -th embedding  $\mathbf{Z}_k^{(l)}$  of  $k$ -cells in a CCNN can be unified as follows:

$$\mathbf{Z}_k^{(l)} = \psi[\oplus(\mathbf{A}_{k-1}^B \mathbf{Z}_{k-1}^{(l-1)} \mathbf{W}_B^{(l-1)}, \mathbf{A}_{k+1}^C \mathbf{Z}_{k+1}^{(l-1)} \mathbf{W}_C^{(l-1)}, \mathbf{A}_k^U \mathbf{Z}_k^{(l-1)} \mathbf{W}_U^{(l-1)}, \mathbf{A}_k^L \mathbf{Z}_k^{(l-1)} \mathbf{W}_L^{(l-1)}, \mathbf{Z}_k^{(l-1)} \mathbf{W}^{(l-1)})], \quad (1)$$

where  $\oplus$  denotes a differentiable, permutation-invariant aggregation function, *e.g.*, sum, mean, or concatenation, and  $\psi$  is a nonlinear updating function such as ReLU. The initial embedding is defined as  $\mathbf{Z}_k^{(0)} = \mathbf{X}_k$ . After  $L$  layers of message passing, the embeddings of 0-cells (nodes), 1-cells (edges), and 2-cells (faces) are denoted as  $\mathbf{Z}_0^{(L)}$ ,  $\mathbf{Z}_1^{(L)}$ , and  $\mathbf{Z}_2^{(L)}$ , respectively. The final representation  $\mathbf{Z}_X$  of the cellular complex  $X$  is then given by:

$$\mathbf{Z}_X = \text{READOUT} \left( \sum_{k=0}^2 \text{MLP}_k(\phi_k(\mathbf{Z}_k^{(L)})) \right), \quad (2)$$



$h_i$ . Note that in practice, we set  $N_m = 3$ , corresponding to maximum ring sizes of 6, 9, and 12, respectively.

**Common Space Contrast** In Common space, all these embeddings are passed through a projection group  $PG^{com}$  and mapped into a common embedding space  $\mathcal{S}$ , where we assume that all granularities offer informative semantic cues. We denote the resulting projected feature of the  $j$ -th view of graph  $G_i$  as  $z_i^{g_j}$ , where the 0-th (GIN-based) view is denoted by  $z_i$ . We then optimize a contrastive objective (Chen et al. 2020; Hassani and Ahmadi 2020; Li et al. 2023) in this space, which encourages consistency across different views of the same graph while enforcing distinction from other graphs. The loss is formally defined as:

$$\mathcal{L}_{COM} = - \sum_i \sum_{\substack{j,k \in [0, N_m] \\ j \neq k}} \log \left[ \frac{s(i, j, k)}{s(i, j, k) + \sum_{\substack{i' \in [1, n], i' \neq i \\ j', k' \in [0, N_m]}} s(i', j', k')} \right] \quad (3)$$

where  $s(i, j, k) = \exp(c(z_i^j, z_i^k)/\tau)$  defines the similarity between a positive pair,  $c(\cdot, \cdot)$  denotes cosine similarity, and  $\tau$  is a temperature scaling factor. The term  $s(i', j', k')$  represents the similarity between negative pairs drawn from different samples.

**Multi-Granularity Decoupled Contrast** Although the common embedding space  $\mathcal{S}$  enables joint contrastive learning across multiple granularities, not all topological views are equally informative, i.e., certain granularities may contain misleading patterns for the downstream task. To address this, we introduce a set of projection heads  $PG^{g_a}$  for each granularity  $g_a$ , which map the multi-view representations into granularity-specific decoupled spaces  $\{\mathcal{S}^{g_a}\}_{a=1}^{N_m}$ . In each decoupled space  $\mathcal{S}^{g_a}$ , we treat the projection  $z_i$  of the original graph  $G_i$  as the anchor. The projected features  $\{z_i^{g_{a'}}\}_{a' \neq a}^{N_m}$ , obtained from other granularities, are considered as positive samples, while representations from different graphs serve as negatives. Notably, we exclude the view  $a' = a$  from being used as a positive, as this space-specific assumption posits that the  $g_a$ -view may not contain reliable complementary information. We then apply contrastive learning in space  $\mathcal{S}^{g_a}$ :

$$\mathcal{L}_{\mathcal{S}^{g_a}} = - \sum_i \sum_{\substack{j=0, k \neq a, \\ k \in [0, N_m]}} \log \left[ \frac{s(i, j, k)}{s(i, j, k) + \sum_{\substack{i' \in [1, n], i' \neq i \\ j', k' \in [0, N_m]}} s(i', j', k')} \right] \quad (4)$$

By aggregating the contrastive losses across all granularity-specific decoupled spaces, we obtain the MGDC objective:

$$\mathcal{L}_{MGDC} = \sum_{a=1}^{N_m} \Omega_{g_a} \cdot \mathcal{L}_{\mathcal{S}^{g_a}}, \quad (5)$$

where  $\Omega_{g_a}$  is the trustworthiness weight of granularity  $g_a$ , reflecting the reliability of its representation space, which is detailed in the following part.

**Granularity-specific Weighting** We quantify this trustworthiness by estimating the uncertainty of the learned representations in each space. Intuitively, if a decoupled space

induces representations with high clustering uncertainty, it is likely that the corresponding granularity fails to capture consistent semantic patterns. To model this, we adopt a Dirichlet distribution-based formulation (Jøssang 2016; Han et al. 2021). For each sample, we first compute the clustering probability  $p_{iq}^{g_a}$  of assigning the feature  $z_i^{g_a}$  to cluster  $q$ , using the Student’s t-distribution:

$$p_{iq}^{g_a} = \frac{(1 + \|z_i^{g_a} - c_q^{g_a}\|^2/\mu)^{-\frac{\mu+1}{2}}}{\sum_{q'} (1 + \|z_i^{g_a} - c_{q'}^{g_a}\|^2/\mu)^{-\frac{\mu+1}{2}}}, \quad (6)$$

where  $\mu$  is a degrees-of-freedom parameter, and  $c_q^{g_a}$  is the centroid of cluster  $q$  in space  $\mathcal{S}^{g_a}$ , formalized as:

$$c_q^{g_a} = \begin{cases} \frac{1}{N_m+1} \sum_{j=0}^{N_m} z_{q,j}, & \text{in common space } \mathcal{S}, \\ \frac{1}{N_m} \sum_{j=0, j \neq a}^{N_m} z_{q,j}, & \text{in decoupled space } \mathcal{S}^{g_a}, \end{cases} \quad (7)$$

We interpret the clustering distribution  $p_q^{g_a}$  as evidence and map it to a Dirichlet distribution with parameters  $d_q^{g_a} = p_q^{g_a} + 1$ . The overall uncertainty mass  $u^{g_a}$  and belief masses  $b_q^{g_a}$  are computed as:

$$b_q^{g_a} = \frac{p_q^{g_a}}{\sum_{q'} d_{q'}^{g_a}}, \quad u^{g_a} = \frac{K}{\sum_{q'} d_{q'}^{g_a}}, \quad (8)$$

where  $K$  is the number of clusters (i.e. batchsize). A lower Dirichlet uncertainty indicates a more confident and compact clustering of features, implying a higher trustworthiness of the corresponding granularity. Finally, we aggregate the instance-wise uncertainty into the trustworthiness score for granularity  $g_a$  as:

$$\Omega_{g_a} = \text{sgm} \left( \sum_{i=1}^n \sum_{q=1}^K \mathbb{I}_{iq} \cdot \left[ \text{dgm}(d_q^{g_a}) - \text{dgm} \left( \sum_{q'} d_{q'}^{g_a} \right) \right] \right), \quad (9)$$

where  $\text{dgm}(\cdot)$  is the digamma function,  $\mathbb{I}_{iq}$  indicates the cluster assignment, and  $\text{sgm}(\cdot)$  is the sigmoid function to normalize the result. This adaptive mechanism enables the model to emphasize the contributions of reliable granularities while suppressing the influence of potentially noisy or uninformative ones.

**Training Objective** To integrate the complementary strengths of both the common and granularity-specific contrastive views, we formulate the overall training objective as a weighted combination of the joint contrastive loss in the common space  $\mathcal{S}$  and the granularity-aware decoupled contrastive losses  $\mathcal{L}_{\mathcal{S}^{g_a}}$ :

$$\mathcal{L}_{\text{all}} = \alpha \cdot \mathcal{L}_{COM} + \frac{\beta}{N_m} \sum_{a \in \{1, \dots, N_m\}} \Omega_{g_a} \cdot \mathcal{L}_{\mathcal{S}^{g_a}}, \quad (10)$$

where  $\alpha$  and  $\beta$  are balancing hyper-parameters that control the relative contributions of common contrast and decoupled contrast. While the contrastive learning framework encourages the model to align and discriminate across multiple topological granularities during training, we further leverage this multi-granularity information at inference time. Specifically, we concatenate the embeddings after the encoders

Dataset	NCI1	PROTEINS	MUTAG	NCI109	IMDB-B	IMDB-M	A.R. ↓
node2vec	54.9 ± 1.6	57.5 ± 3.6	72.6 ± 10.0	-	-	-	13.0
sub2vec	52.8 ± 1.5	53.0 ± 5.6	61.1 ± 15.8	-	55.3 ± 1.5	-	13.8
graph2vec	73.2 ± 1.8	73.3 ± 2.0	83.2 ± 9.3	-	71.1 ± 0.5	-	11.0
InfoGraph	76.2 ± 1.0	74.4 ± 0.3	89.0 ± 1.1	76.2 ± 1.3	73.0 ± 0.9	48.1 ± 0.3	7.5
GraphCL	77.9 ± 0.4	74.4 ± 0.5	86.8 ± 1.3	78.1 ± 0.4	71.2 ± 0.4	48.9 ± 0.3	8.0
ADGCL	73.9 ± 0.8	73.3 ± 0.5	88.7 ± 1.9	72.4 ± 0.4	70.2 ± 0.7	48.1 ± 0.4	10.2
JOAO	78.1 ± 0.5	74.6 ± 0.4	87.4 ± 1.0	77.2 ± 0.6	70.2 ± 3.1	48.9 ± 1.2	8.5
JOAOv2	78.4 ± 0.5	74.1 ± 1.1	87.7 ± 0.8	78.2 ± 0.8	70.8 ± 0.3	49.2 ± 0.9	7.0
RGCL	78.1 ± 1.0	75.0 ± 0.4	87.7 ± 1.0	77.7 ± 0.3	71.9 ± 0.9	49.3 ± 0.4	6.2
SimGRACE	79.1 ± 0.4	75.3 ± 0.1	89.0 ± 1.3	78.4 ± 0.4	71.3 ± 0.8	49.1 ± 0.8	4.3
HTML	78.2 ± 0.7	75.0 ± 0.3	88.9 ± 0.8	77.9 ± 0.2	71.7 ± 0.4	48.9 ± 0.6	6.0
DRGCL	78.7 ± 0.4	75.2 ± 0.6	89.5 ± 0.6	77.8 ± 0.4	72.0 ± 0.5	49.5 ± 0.5	4.2
CellCLAT	79.4 ± 0.2	75.7 ± 0.1	89.7 ± 0.3	78.9 ± 0.4	73.4 ± 0.1	50.6 ± 0.2	<u>2.0</u>
HTG-GCL	<b>80.7 ± 0.4</b> ***	<b>76.2 ± 0.5</b> **	<b>90.4 ± 0.5</b> **	<b>79.7 ± 0.4</b> ***	<b>74.0 ± 0.4</b> ***	<b>51.5 ± 0.5</b> ***	<b>1.0</b>

Table 1: Unsupervised representation learning classification accuracy (%) on TU datasets. A.R denotes the average rank of the results. The best results are highlighted in **bold**, and the second best results are highlighted with underline. Statistical significance is determined via a one-sided two-sample t-test. \*, \*\*, and \*\*\* indicate p-values less than 0.1, 0.05, and 0.01, denoting marginally, statistically, and strongly significant results, respectively.

Dataset	NCI1	PROTEINS	NCI109	IMDB-B	IMDB-M	A.R. ↓
InfoGraph	68.9 ± 0.8	69.5 ± 0.9	67.8 ± 0.7	68.5 ± 0.6	43.8 ± 0.2	8.2
GraphCL	69.6 ± 0.3	69.7 ± 0.1	69.2 ± 0.3	67.8 ± 1.3	43.9 ± 0.3	6.0
ADGCL	65.6 ± 0.6	68.5 ± 2.0	65.7 ± 0.8	67.7 ± 0.9	41.9 ± 0.9	10.8
JOAO	69.1 ± 0.1	70.9 ± 2.0	69.2 ± 0.3	67.6 ± 0.8	43.9 ± 1.0	6.6
JOAOv2	69.1 ± 0.4	70.1 ± 1.4	69.2 ± 0.2	68.3 ± 1.0	43.6 ± 1.1	6.8
RGCL	70.2 ± 0.7	71.2 ± 0.9	69.1 ± 0.3	68.9 ± 0.3	44.5 ± 0.9	3.4
SimGRACE	69.3 ± 0.3	71.3 ± 0.7	69.0 ± 0.2	68.6 ± 0.7	44.2 ± 0.6	4.6
HTML	69.4 ± 0.1	69.1 ± 1.5	68.9 ± 0.1	68.4 ± 0.3	43.2 ± 0.3	7.8
DRGCL	69.3 ± 0.2	70.3 ± 0.4	68.4 ± 0.2	68.0 ± 0.6	44.3 ± 0.5	6.4
CellCLAT	70.4 ± 0.5	71.8 ± 0.1	69.5 ± 0.3	68.5 ± 0.4	44.0 ± 0.3	<u>3.0</u>
HTG-GCL	<b>70.8 ± 0.6</b> *	<b>72.3 ± 0.4</b> **	<b>70.6 ± 0.4</b> ***	<b>69.1 ± 0.5</b> *	<b>45.2 ± 0.7</b> *	<b>1.0</b>

Table 2: Semi-supervised representation learning classification accuracy (%) on TU datasets.

from all granularity-specific views  $\{h_i^{g_a}\}_{a \in \{1, \dots, N_m\}}$  and the graph representation  $h_i$  as the final graph-level embedding:

$$h_i^{\text{final}} = \text{CONCAT}(h_i, h_i^{g_1}, h_i^{g_2}, \dots, h_i^{g_{N_m}}). \quad (11)$$

$h_i^{\text{final}}$  is then used for downstream classification, allowing the model to fully exploit the hierarchical topological granularity encoded by the cellular complex transformations.

## 5 Experiments

### 5.1 Experimental Setup

**Benchmarks** We evaluate the performance of HTG-GCL on six widely adopted benchmarks from the TU collection (Morris et al. 2020), including four bioinformatics datasets (NCI1, PROTEINS, MUTAG, NCI109) and two social network datasets (IMDB-B, IMDB-M). For unsupervised evaluation, all datasets are used, while in the semi-supervised setting, MUTAG is excluded due to its small size and severe class imbalance under k-fold cross-validation.

In the unsupervised setting, HTG-GCL is trained on the full dataset without label supervision. The learned graph-level representations are evaluated via a downstream classification task using an SVM classifier with 10-fold cross-validation. We report the mean and standard deviation of classification accuracy across five different random seeds.

For the semi-supervised evaluation, the same training procedure is used, but only 10% of the labeled data is employed for SVM training. To leverage the remaining unlabeled data, we adopt a pseudo-labeling strategy: a classifier is first trained on the small labeled subset and then used to generate pseudo-labels for the unlabeled samples. A final model is trained using both labeled and pseudo-labeled data.

**Compared Baselines** In the unsupervised scenario, we compare HTG-GCL against fourteen representative baselines: Node2Vec (Grover and Leskovec 2016), Sub2Vec (Adhikari et al. 2018), Graph2Vec (Narayanan et al. 2017), InfoGraph (Sun et al. 2020), GraphCL (You et al. 2020), ADGCL (Suresh et al. 2021), JOAO (You et al. 2021),



Dataset	NCI1	MUTAG	IMDB-B	IMDB-M
One granularity (ring 6)	79.3 $\pm$ 0.3	89.7 $\pm$ 0.3	73.2 $\pm$ 0.3	50.5 $\pm$ 0.3
One granularity (ring 9)	79.5 $\pm$ 0.4	89.3 $\pm$ 0.5	73.3 $\pm$ 0.2	50.8 $\pm$ 0.6
One granularity (ring 12)	79.5 $\pm$ 0.3	89.4 $\pm$ 0.4	72.8 $\pm$ 0.5	50.7 $\pm$ 0.5
Two granularities (ring 6, 9)	79.9 $\pm$ 0.4	89.9 $\pm$ 0.3	73.4 $\pm$ 0.2	51.0 $\pm$ 0.3
Two granularities (ring 6,12)	80.1 $\pm$ 0.6	89.5 $\pm$ 0.5	73.2 $\pm$ 0.4	51.1 $\pm$ 0.4
Two granularities (ring 9, 12)	80.1 $\pm$ 0.4	89.7 $\pm$ 0.6	73.4 $\pm$ 0.3	51.1 $\pm$ 0.3
Two granularities (ring 6, 9) w MGDC	80.3 $\pm$ 0.2	90.1 $\pm$ 0.3	73.7 $\pm$ 0.2	51.2 $\pm$ 0.3
Two granularities (ring 6,12) w MGDC	80.4 $\pm$ 0.4	89.8 $\pm$ 0.4	73.5 $\pm$ 0.2	51.4 $\pm$ 0.4
Two granularities (ring 9, 12) w MGDC	79.8 $\pm$ 0.4	89.6 $\pm$ 0.3	73.7 $\pm$ 0.3	51.4 $\pm$ 0.3
HTG-GCL w/o MGDC	79.7 $\pm$ 0.2	89.8 $\pm$ 0.4	73.4 $\pm$ 0.3	50.9 $\pm$ 0.3
HTG-GCL	<b>80.7 <math>\pm</math> 0.4</b>	<b>90.4 <math>\pm</math> 0.5</b>	<b>74.0 <math>\pm</math> 0.4</b>	<b>51.5 <math>\pm</math> 0.5</b>

Table 3: Ablation studies on the unsupervised settings. w and w/o denote with and without, respectively.

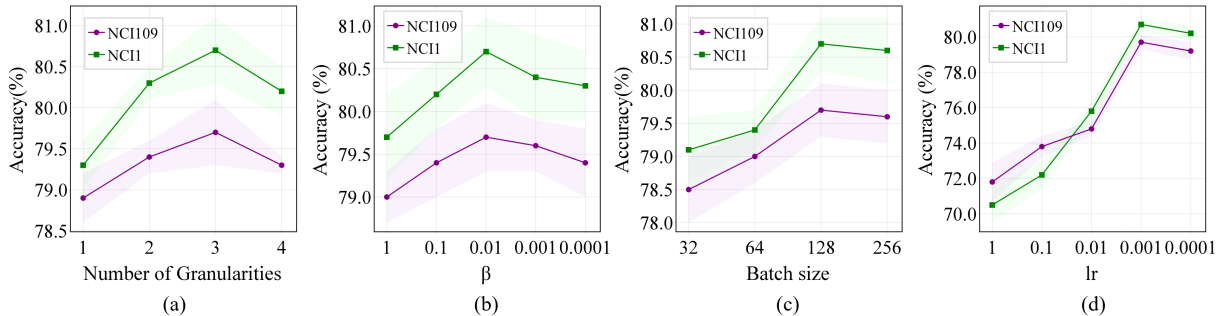


Figure 4: Hyper-parameter analysis.

RGCL (Li et al. 2022), SimGRACE (Xia et al. 2022), HTML (Li et al. 2024), DRGCL (Ji et al. 2024) and CellCLAT (Qin et al. 2025). For the semi-supervised setting, we select ten of these unsupervised baselines for comparison.

**Running Environment** Our method is implemented using PyTorch 1.7. All experiments are conducted on a NVIDIA Tesla V100 GPU, running on Ubuntu 20.04.

## 5.2 Unsupervised Learning

The results of unsupervised graph-level representation learning for downstream graph classification tasks are summarized in Table 1. HTG-GCL consistently achieves the best performance across all datasets, with an impressive average rank of 1.0, outperforming all other methods. This highlights the effectiveness of HTG-GCL in learning discriminative and stable graph representations. Furthermore, from the perspective of statistical significance, HTG-GCL achieves strongly or statistically significant improvements ( $p < 0.01$  or  $p < 0.05$ ) over the second-best method on most datasets, as indicated by the asterisks in the table. This demonstrates that the performance gains of HTG-GCL are not only consistent but also statistically reliable.

## 5.3 Semi-supervised Learning

Table 2 presents the results of semi-supervised graph-level representation learning for downstream classification tasks.

Our method, HTG-GCL, achieves the best overall performance with an average rank of 1.0 across all datasets, outperforming all other methods. This outstanding performance highlights the effectiveness of HTG-GCL in generating high-quality representations during unsupervised pretraining, which can be effectively exploited by downstream classifiers even in low-label regimes through pseudo-labeling.

## 5.4 Ablation Studies

We conducted ablation studies on different granularities and the MGDC module in unsupervised settings, as shown in Table 3. The results clearly indicate that multi-granularity integration outperforms single-granularity topological information, demonstrating the effectiveness of incorporating hierarchical topological information. Moreover, MGDC shows larger improvement in multi-granularity scenarios, highlighting its ability to reduce redundancy and enhance task-relevant information. HTG-GCL demonstrates outstanding performance across all settings, validating the importance of hierarchical topological granularity integration in improving graph contrastive learning tasks.

## 5.5 Hyper-parameter Analysis

**Number of Granularities** We investigate the impact of the number of granularities from 1 to 4 on model performance. The granularities represent the following combinations of granularities: [6], [6,9], [6,9,12], [6,9,12,15]. Figure 4(a) shows the classification accuracy on the NCI1 and

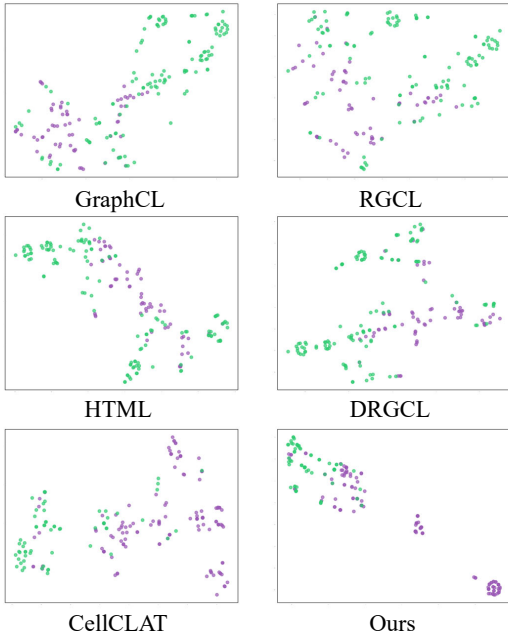


Figure 5: t-SNE visualization of six methods on MUTAG.

NCI109 datasets. The highest accuracy is achieved with 3 granularities, and performance declines when the number of granularities is either lower or higher. The worse performance of more granularities [6,9,12,15] may be due to the scarcity of rings larger than 12 in the NCI dataset.

**Coefficient  $\beta$  for decoupled granularity loss** Figure 4(b) illustrates the classification accuracy on NCI1 and NCI109 dataset across different values of  $\beta \in \{1, 0.1, 0.01, 0.001, 0.0001\}$ . We observe that both datasets achieve the best performance when  $\beta$  is set to 0.01.

**Batch Size** Figure 4(c) shows the classification accuracy after training for 20 epochs using different batch sizes ranging from 32 to 256 on NCI1 and NCI109 datasets. We observe the performance improves as the batch size increases, with the best results achieved at a batch size of 128.

**lr** Figure 4(d) illustrates the sensitivity of the classification performance to the learning rate, with values ranging from 1 to 0.0001. Both datasets achieve the highest accuracy when the learning rate is set to 0.001.

## 5.6 Visualization Results

In Figure 5, we present the t-SNE visualizations of six different methods on the MUTAG dataset. Our method exhibits better separation between classes, indicating its superior ability to learn discriminative representations compared to the baseline methods.

## 5.7 Time Complexity Analysis

**Multi-granularity Graph-to-Complex Transformation** For each maximum ring size  $m \in \{6, 9, 12\}$ , we attach 2-cells to all induced cycles of length at most  $m$ , thereby

Dataset	NCI1	PROTEINS	MUTAG	IMDB-B
Time	43s	14s	53s	151s

Table 4: Transformation time across different datasets.

Dataset	GraphCL	DRGCL	CellCLAT	Ours
PROTEINS	2s	4s	6s	7s
NCI1	17s	18s	15s	18s

Table 5: Training time per epoch across different methods.

constructing a cellular complex  $\{X_m(G_i)\}_m$ . This transformation is a one-time preprocessing step before training. Its complexity is bounded by  $\Theta((|E| + |V|N_m) \text{polylog}|V|)$ , where  $N_m$  is the number of induced cycles with length at most  $m$  (Ferreira et al. 2014). Since HTG-GCL constructs  $N_m = 3$  granularities in practice, the preprocessing cost grows with the total number of rings considered across granularities (thus positively correlated with  $N_m$ ).

**Multi-scale Topological Granularity Decoupled Contrastive Learning** *Message Passing:* According to (Bodnar et al. 2021a), the computational complexity of the message passing scheme is  $\Theta(\sum_{p=1}^d S_p)$ , where  $B_p$  is the maximum boundary size of a  $p$ -cell in  $X$  and  $S_p$  denotes the number of  $p$ -cells. Since HTG-GCL processes  $N_m$  granularities, the passing cost for a batch of  $n$  samples is  $\Theta(N_m n \sum_{p=1}^d S_p)$ . *Contrastive Loss Computation:* HTG-GCL performs contrastive learning in both the common space and granularity-specific decoupled spaces. For a batch of  $n$  samples and embedding dimension  $d$ , the similarity matrix has size  $(N_m n) \times (N_m n)$ , yielding a per-batch complexity  $\Theta(N_m^2 n^2 d)$ . *Uncertainty-based Weighting Module:* In each granularity-specific space, we compute clustering distributions  $p_{iq}^{g_a}$  using Student’s  $t$ -distribution (Eq. 6) and then map them into Dirichlet parameters (Eq. 8-9). For a batch of  $n$  samples with embedding dimension  $d$  and  $K$  clusters ( $K = n$ ), the per-batch complexity is  $\Theta((N_m + 1) n^2 d)$ . Combining the above three parts, the total complexity of HTG-GCL is  $\Theta(N_m n \sum_{p=1}^d S_p) + \Theta(N_m^2 n^2 d) + \Theta((N_m + 1) n^2 d)$ .

## 6 Conclusion

In this paper, we propose HTG-GCL, a framework that addresses the limitations of fixed-granularity contrastive learning. By transforming graphs into multi-scale ring-based cellular complexes, HTG-GCL constructs diverse topological views to capture fine- and coarse-grained structures. To mitigate conflicts across views, we introduce a decoupled contrastive strategy and assign trustworthiness-based weights to granularity-specific representation spaces. Extensive experiments on benchmark datasets show that HTG-GCL achieves consistent improvements over state-of-the-art methods. This work highlights the importance of hierarchical granularity modeling and opens new directions for adaptive contrastive learning on topological graph structures.



## Acknowledgements

The authors would like to thank the editors and reviewers for their valuable comments. This work is supported by the National Natural Science Foundation of China No. 62406313, 2023 Special Research Assistant Grant Project of the Chinese Academy of Sciences.

## References

- Adhikari, B.; Zhang, Y.; Ramakrishnan, N.; and Prakash, B. A. 2018. Sub2vec: Feature learning for subgraphs. In *Advances in Knowledge Discovery and Data Mining: 22nd Pacific-Asia Conference, PAKDD 2018, Melbourne, VIC, Australia, June 3-6, 2018, Proceedings, Part II* 22, 170–182. Springer.
- Bodnar, C.; Frasca, F.; Otter, N.; Wang, Y.; Lio, P.; Montufar, G. F.; and Bronstein, M. 2021a. Weisfeiler and lehmango cellular: Cw networks. *Advances in neural information processing systems*, 34: 2625–2640.
- Bodnar, C.; Frasca, F.; Wang, Y.; Otter, N.; Montufar, G. F.; Lio, P.; and Bronstein, M. 2021b. Weisfeiler and lehmango topological: Message passing simplicial networks. In *International Conference on Machine Learning*, 1026–1037. PMLR.
- Chen, T.; Kornblith, S.; Norouzi, M.; and Hinton, G. E. 2020. A Simple Framework for Contrastive Learning of Visual Representations. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, 1597–1607. PMLR.
- Ferreira, R.; Grossi, R.; Rizzi, R.; Sacomoto, G.; and Sagot, M.-F. 2014. Amortized-delay algorithm for listing chordless cycles in undirected graphs. In *European Symposium on Algorithms*, 418–429. Springer.
- Gao, H.; Li, J.; Qiang, W.; Si, L.; Sun, F.; and Zheng, C. 2022. Bootstrapping Informative Graph Augmentation via A Meta Learning Approach. In Raedt, L. D., ed., *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, 3001–3007. ijcai.org.
- Grover, A.; and Leskovec, J. 2016. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, 855–864.
- Hajij, M.; Istvan, K.; and Zamzmi, G. 2020. Cell complex neural networks. *arXiv preprint arXiv:2010.00743*.
- Han, Z.; Zhang, C.; Fu, H.; and Zhou, J. T. 2021. Trusted Multi-View Classification. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net.
- Hansen, J.; and Ghrist, R. 2019. Toward a spectral theory of cellular sheaves. *Journal of Applied and Computational Topology*, 3(4): 315–358.
- Hassani, K.; and Ahmadi, A. H. K. 2020. Contrastive Multi-View Representation Learning on Graphs. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, volume 119 of *Proceedings of Machine Learning Research*, 4116–4126. PMLR.
- 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 2024, February 20-27, 2024, Vancouver, Canada*, 12810–12820. AAAI Press.
- Jøsang, A. 2016. *Subjective Logic - A Formalism for Reasoning Under Uncertainty*. Artificial Intelligence: Foundations, Theory, and Algorithms. Springer. ISBN 978-3-319-42335-7.
- Ju, W.; Wang, Y.; Qin, Y.; Mao, Z.; Xiao, Z.; Luo, J.; Yang, J.; Gu, Y.; Wang, D.; Long, Q.; Yi, S.; Luo, X.; and Zhang, M. 2024. Towards Graph Contrastive Learning: A Survey and Beyond. *arXiv:2405.11868*.
- Li, J.; Jin, Y.; Gao, H.; Qiang, W.; Zheng, C.; and Sun, F. 2024. Hierarchical Topology Isomorphism Expertise Embedded Graph Contrastive Learning. 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 2024, February 20-27, 2024, Vancouver, Canada*, 13518–13527. AAAI Press.
- Li, J.; Qiang, W.; Zheng, C.; Su, B.; Razzak, F.; Wen, J.; and Xiong, H. 2023. Modeling Multiple Views via Implicitly Preserving Global Consistency and Local Complementarity. *IEEE Trans. Knowl. Data Eng.*, 35(7): 7220–7238.
- Li, S.; Wang, X.; Zhang, A.; Wu, Y.; He, X.; and Chua, T.-S. 2022. Let invariant rationale discovery inspire graph contrastive learning. In *International conference on machine learning*, 13052–13065. PMLR.
- Madhu, H.; and Chepuri, S. P. 2023. TopoSRL: topology preserving self-supervised simplicial representation learning. *Advances in Neural Information Processing Systems*, 36.
- Morris, C.; Kriege, N. M.; Bause, F.; Kersting, K.; Mutzel, P.; and Neumann, M. 2020. TUDataset: A collection of benchmark datasets for learning with graphs. *arXiv preprint arXiv:2007.08663*.
- Narayanan, A.; Chandramohan, M.; Venkatesan, R.; Chen, L.; Liu, Y.; and Jaiswal, S. 2017. graph2vec: Learning distributed representations of graphs. *arXiv preprint arXiv:1707.05005*.
- Papillon, M.; Sanborn, S.; Hajij, M.; and Miolane, N. 2024. Architectures of Topological Deep Learning: A Survey of Message-Passing Topological Neural Networks. *arXiv:2304.10031*.
- Qin, B.; Ji, Q.; Li, J.; Wang, Y.; Wu, X.; Cao, J.; and Xu, F. 2025. CellCLAT: Preserving Topology and Trimming Redundancy in Self-Supervised Cellular Contrastive Learning. *arXiv:2505.21587*.

- Sun, F.-Y.; Hoffman, J.; Verma, V.; and Tang, J. 2020. InfoGraph: Unsupervised and Semi-supervised Graph-Level Representation Learning via Mutual Information Maximization. In *International Conference on Learning Representations*. OpenReview. net.
- Suresh, S.; Li, P.; Hao, C.; and Neville, J. 2021. Adversarial graph augmentation to improve graph contrastive learning. *Advances in Neural Information Processing Systems*, 34: 15920–15933.
- 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 *Proceedings of the ACM web conference 2022*, 1070–1079.
- Xu, K.; Hu, W.; Leskovec, J.; and Jegelka, S. 2018. How powerful are graph neural networks? *arXiv preprint arXiv:1810.00826*.
- Yang, M.; Isufi, E.; and Leus, G. 2022. Simplicial Convolutional Neural Networks. In *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2022, Virtual and Singapore, 23-27 May 2022*, 8847–8851. IEEE.
- You, Y.; Chen, T.; Shen, Y.; and Wang, Z. 2021. Graph contrastive learning automated. In *International conference on machine learning*, 12121–12132. PMLR.
- You, Y.; Chen, T.; Sui, Y.; Chen, T.; Wang, Z.; and Shen, Y. 2020. Graph contrastive learning with augmentations. *Advances in neural information processing systems*, 33: 5812–5823.
- Zhao, X.; Ying, C.; and Yu, T. 2024. Graph Learning with Distributional Edge Layouts. *arXiv:2402.16402*.