CLHYPFORMER: TOPOLOGY-AWARE HYPERBOLIC HYPERGRAPH TRANSFORMER FOR GRAPH REPRESENTATION LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Learning effective representations from complex graphs requires simultaneously capturing hierarchical geometry and higher-order topology. However, conventional Euclidean embeddings can distort hierarchical relationships, whereas existing hypergraph construction strategies based on Euclidean proximity or random walks may introduce structurally inconsistent connections. We propose CLhypformer, a topology-aware hyperbolic hypergraph Transformer that jointly addresses these geometric and topological limitations. CLhypformer first constructs hyperedges through hyperbolic K-nearest-neighbor relationships, enabling topology construction directly in a geometry that better accommodates hierarchical structure. On top of the resulting hypergraph, we develop a topology-aware hyperbolic contrastive learning framework that generates asymmetric structural views through a topology-aware discriminator and employs a debiased three-level contrastive objective to align representations across multiple structural granularities. The resulting geometry–topology coupling allows the model to exploit hierarchical proximity while suppressing structurally inconsistent perturbations and spurious connections during representation learning. Extensive experiments on multiple graph benchmarks demonstrate that CLhypformer achieves state-of-the-art or highly competitive performance for node classification. Comprehensive ablation and robustness analyses further show that both hyperbolic topology construction and topology-aware contrastive alignment contribute substantially to the overall performance, with their combination yielding consistent improvements across diverse graph structures. These results demonstrate the potential of jointly modeling non-Euclidean geometry and higher-order topology for robust graph representation learning.

1 INTRODUCTION

Node classification is a fundamental task in graph representation learning, with broad applications in citation networks, social networks, bioinformatics, and recommendation systems (Yi et al., 2022). Real-world graphs, however, often exhibit complex structural patterns that extend beyond local pairwise relationships, including hierarchical organization and higher-order interactions. Such structures pose fundamental challenges to conventional graph representation learning methods. Hypergraph neural networks have therefore attracted increasing attention by explicitly modeling interactions among groups of nodes rather than restricting information propagation to pairwise edges. Despite their ability to capture higher-order dependencies, most existing hypergraph learning methods represent nodes and hypergraph structures in Euclidean spaces (Antelmi et al., 2023). This geometric assumption can be limiting for hierarchical data, because the volume of Euclidean space grows polynomially with radius, whereas the number of nodes in hierarchical structures can increase much more rapidly with depth. Consequently, Euclidean embeddings may require substantial metric distortion to preserve hierarchical relationships (Chen et al., 2026).

Hyperbolic representation learning provides a natural alternative for modeling such non-Euclidean structures. By embedding representations in negatively curved Riemannian manifolds, hyperbolic geometry offers exponentially expanding capacity with increasing radius and can therefore represent tree-like and hierarchical structures with substantially lower distortion than Euclidean spaces

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.

(Peng et al., 2022). However, exploiting the advantages of hyperbolic geometry for higher-order graph representation learning remains challenging. In particular, the geometric space in which node representations are learned should be consistent with the topology used to define higher-order interactions. Existing hypergraph construction strategies are predominantly based on Euclidean K-nearest-neighbor (KNN) relationships or random-walk procedures. When applied to graphs with hierarchical structure, these approaches may construct hyperedges according to geometrically inappropriate neighborhoods, thereby introducing spurious connections and structurally inconsistent higher-order relationships. Such topology–geometry mismatch can propagate noisy information through the hypergraph and ultimately compromise the quality of learned representations. This motivates a fundamental question: how can hyperbolic geometry and higher-order topology be jointly modeled so that hypergraph construction and representation learning remain structurally consistent?

To address this challenge, we propose CLhypformer, a topology-aware hyperbolic hypergraph Transformer for graph representation learning.

The main contributions of this work are summarized as follows:

(1) Topology-aware hyperbolic contrastive learning. We develop a hypergraph contrastive learning framework specifically designed for hyperbolic representations. A dynamic topology-aware discriminator is employed to generate asymmetric structural views, while a debiased three-level contrastive objective promotes representation alignment at multiple structural granularities. By incorporating topology into view generation and contrastive optimization, the proposed framework reduces the structural inconsistency and false-negative effects that can arise from generic contrastive augmentation strategies (Li et al., 2025; Chuang et al., 2020).

(2) Hyperbolic KNN-based hypergraph construction. We introduce an undirected hypergraph construction mechanism based on KNN relationships in a Lorentz manifold with learnable curvature. Unlike conventional Euclidean-space neighborhood construction, the proposed strategy determines higher-order neighborhoods according to hyperbolic distances, thereby reducing geometric distortion when modeling hierarchical structures and mitigating noisy connections associated with conventional topology construction strategies (Yang et al., 2022; Nickel & Kiela, 2018).

(3) Comprehensive empirical evaluation. We conduct extensive experiments and ablation studies across graph benchmarks to evaluate the effectiveness and robustness of CLhypformer. The results demonstrate competitive or superior performance for node classification under diverse structural conditions. In particular, robustness experiments under structural masking and controlled ablations show that topology-aware hyperbolic contrastive alignment provides a substantial contribution to the overall performance, highlighting the importance of jointly modeling geometry and topology for robust graph representation learning.

2 PRELIMINARIES

2.1 LORENTZ MODEL OF HYPERBOLIC SPACE

Lorentz Model. The Lorentz model represents a d-dimensional hyperbolic space as a submanifold of a $(d+1)$-dimensional pseudo-Euclidean space (Nickel & Kiela, 2018; Law et al., 2019). Given a positive curvature-scale parameter c, the Lorentz manifold is defined as:

$$\{x \in \mathbb{R}^{d+1} | \langle x, x \rangle_{\mathcal{L}} = -\frac{1}{c}, x_0 > 0\}, \tag{1}$$

where x_0 denotes the time-like component, $x_{1:d}$ represents the spatial components, and $\langle \cdot, \cdot \rangle_{\mathcal{L}}$ is the Lorentz inner product. The resulting manifold has constant sectional curvature $-c$, where c controls the geometric scale of the hyperbolic space.

Adaptive Curvature. To accommodate datasets with different intrinsic hierarchical structures, we treat the curvature scale c as a learnable parameter rather than fixing it a priori. Given an initial Euclidean representation $h \in \mathbb{R}^d$, it is first mapped to the Lorentz manifold using a curvature-dependent initialization:

$$c = Softplus(\widetilde{c}), \tag{2}$$

$$\exp_o^c(h), \tag{3}$$

where $\widetilde{c}$ is an unconstrained learnable scalar and o denotes the origin of the corresponding tangent space. This adaptive parameterization allows the geometry to adjust to the structural characteristics of the input graph while ensuring that the resulting representations satisfy the manifold constraint (Fu et al., 2021; Chami et al., 2019).

Lorentzian Distance. Unlike Euclidean distances, which measure proximity along straight lines in a flat space, hyperbolic distances account for the intrinsic geometry of negatively curved manifolds. For two points $x, y \in \mathbb{R}^{d,c}$, their geodesic distance is given by

$$d_{\mathcal{L}}(x, y) = \frac{1}{\sqrt{c}} arcosh(-c \langle x, y \rangle_{\mathcal{L}}). \tag{4}$$

Accordingly, the Lorentzian inner product provides the geometric basis for measuring similarity and neighborhood relationships in hyperbolic space:

$$\langle x, y \rangle_{\mathcal{L}} = -x_0 y_0 + \sum_{i=1}^{d} x_i y_i. \tag{5}$$

Exponential and Logarithmic Mappings. Because Euclidean vector operations are not intrinsically defined on a curved manifold, intermediate computations can be performed in a tangent space and subsequently mapped back to the manifold. The exponential and logarithmic maps establish this bidirectional correspondence between a tangent space and the Lorentz manifold (Chami et al., 2019):

$$\exp_x^c : T_x H^{d,c} \to H^{d,c}, \tag{6}$$

$$\log_x^c : H^{d,c} \to T_x H^{d,c}. \tag{7}$$

Here, $T_x H^{d,c}$ denotes the tangent space at x.

2.2 Hyperbolic Layers

Standard neural operators, including linear transformations, nonlinear activations, and normalization, are intrinsically defined in Euclidean vector spaces and therefore cannot be directly applied to Lorentz-manifold representations without accounting for the underlying geometric constraints (Yang et al., 2024). Following the tangent-space formulation, a hyperbolic layer first maps the input representation to the tangent space, applies the corresponding Euclidean operator, and then maps the transformed representation back to the Lorentz manifold.

For a linear transformation parameterized by a weight matrix W and bias b, the hyperbolic counterpart can be expressed as

$$\exp_x^c(W \log_o^c(x) + b) \tag{8}$$

where o denotes the origin of the tangent space, and W and b are learnable parameters. Similarly, nonlinear transformations and normalization operations can be implemented in the tangent space before mapping the resulting representations back to $H^{d,c}$.

3 METHODOLOGY

3.1 OVERALL ARCHITECTURE AND FORWARD PROPAGATION

CLhypformer couples representation geometry with higher-order topology in a unified forward process. As shown in Figure 1, Euclidean node features are first mapped to the Lorentz manifold, where hyperbolic KNN constructs the initial hypergraph. A topology-aware discriminator then produces asymmetric structural views for multi-level contrastive alignment, while the hyperbolic hypergraph propagation module aggregates higher-order information and outputs node-classification predictions.

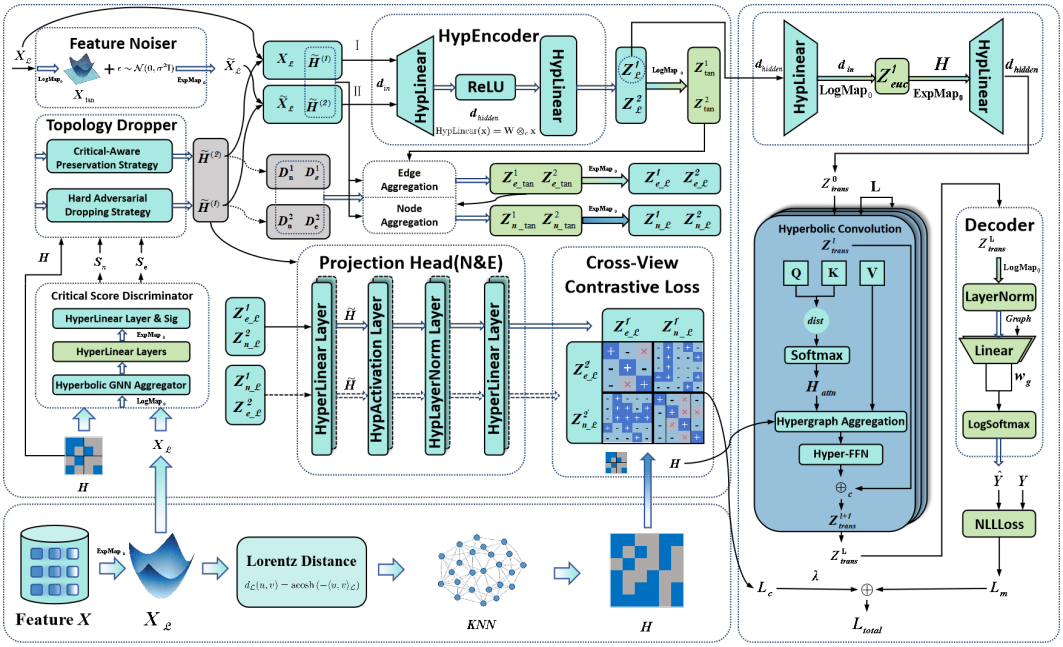


Figure 1: Overall architecture of CLhypformer. The framework consists of three main components: (1) Topology Construction, which maps node features to the Lorentz manifold and constructs a higher-order hypergraph using hyperbolic KNN; (2) Topology-Aware Contrastive Learning, which generates asymmetric structural views through criticality-aware retention and hard perturbation, followed by cross-view alignment at multiple representation levels; and (3) Hyperbolic Hypergraph Propagation, which performs hyperbolic self-attention and higher-order hypergraph aggregation for downstream node classification.

Hyperbolic Geometry Mapping and Data Initialization

Given an input hypergraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ and $\mathcal{E}$ denote the node and hyperedge sets, respectively, we represent its topology using an incidence matrix $H \in \{0,1\}^{|\mathcal{V}|\times|\mathcal{E}|}$ and its node attributes using a Euclidean feature matrix $X \in \mathbb{R}^{|\mathcal{V}|\times d_{\text{in}}}$ (Feng et al., 2019). To obtain a representation space suitable for hierarchical structures, the Euclidean node features are mapped to the Lorentz manifold $\mathcal{L}^d$ with learnable curvature $\kappa > 0$. Specifically, using the exponential map at the origin o, the initial hyperbolic representations are given by

$$\mathbf{X}_{\mathcal{L}} = exp_o^k(\mathbf{X}) = \left(\frac{1}{\sqrt{\kappa}}\cosh(\sqrt{\kappa}\parallel \mathbf{X} \parallel), \frac{1}{\parallel \mathbf{X} \parallel}\sinh(\sqrt{\kappa}\parallel \mathbf{X} \parallel)\mathbf{X}\right). \tag{9}$$

Topology-Aware Discrimination and Asymmetric Dual-View Generation

Generic graph perturbation strategies may remove structurally informative connections without considering their roles in the underlying topology. To generate structurally meaningful contrastive views, CLhypformer employs a topology-aware discriminator to estimate the structural criticality

of node–hyperedge associations. Specifically, the manifold representations are first mapped to the tangent space through the logarithmic map, where a lightweight hypergraph message-passing network extracts local structural dependencies. The discriminator then predicts a criticality matrix

$$\mathbf{W}_{\text{crit}} = \sigma\left(\text{HyperGNN}_{\text{disc}}\left(\mathbf{X}_{\mathcal{L}}, \mathbf{H}\right)\right), \tag{10}$$

where $\sigma(\cdot)$ denotes the sigmoid function and $W_{\text{crit}} \in [0,1]^{\{|\mathcal{V}|\times|\mathcal{E}|}$ quantifies the estimated structural importance of each node–hyperedge association.

Based on $\mathbf{W}_{\text{crit}}$, we construct two complementary views with different structural perturbation patterns. The first is a structure-preserving view $\mathcal{G}^{(1)}$, which preferentially retains associations with high structural importance. Its corresponding mask matrix $M^{(1)}$ is sampled as

$$\mathbf{M}_{i,j}^{(1)} \sim Bernoulli(1 - p_{\text{base}} \cdot (1 - \mathbf{W}_{\text{crit,ij}})). \tag{11}$$

The second is a hard-perturbed view $\mathcal{G}^{(2)}$, which assigns larger deletion probabilities to structurally important associations, subject to a maximum perturbation probability $p_{\max} = 0.8$, producing the mask matrix $M^{(2)}$:

$$\mathbf{M}_{i,j}^{(2)} \sim Bernoulli(1 - min(p_{\max}, p_{\text{base}} \cdot \mathbf{W}_{\text{crit,ij}})). \tag{12}$$

Topology-Aware Debiased Three-Level Contrastive Learning

The two views are subsequently processed by the hyperbolic encoder to obtain node- and hyperedge-level representations, denoted by $Z_v^{(1)}, Z_v^{(2)}$ and $Z_e^{(1)}, Z_e^{(2)}$, respectively (Li et al., 2025; Lee & Shin, 2023). For two manifold representations u and v, their geometric discrepancy is measured using the Lorentz geodesic distance

$$d_{\mathcal{L}}(\mathbf{u}, \mathbf{v}) = \frac{1}{\sqrt{\kappa}} \text{arcosh}\left(-\kappa \langle \mathbf{u}, \mathbf{v} \rangle_{\mathcal{L}}\right). \tag{13}$$

To mitigate false-negative effects caused by structurally related samples being treated as negatives, we construct a topology-aware mask M_{topo} based on multi-hop reachability and incorporate a class-prior debiasing term into the contrastive objective. At the node level, the resulting contrastive loss is formulated as

$$\mathcal{L}_{\text{node}} = -log \frac{\exp\left(-d_{\mathcal{L}}(\mathbf{z}_{v,i}^{(1)}, \mathbf{z}_{v,i}^{(2)})/\tau\right)}{\sum_{j \neq i} \mathbf{M}_{\text{topo,ij}} \cdot exp\left(-d_{\mathcal{L}}(\mathbf{z}_{v,i}^{(1)}, \mathbf{z}_{v,j}^{(2)})/\tau\right) + Debias(\mathbf{z}_{v,i}^{(1)})}, \tag{14}$$

where τ denotes the temperature parameter. The overall contrastive objective combines node-level, hyperedge-level, and cross-view alignment:

$$\mathcal{L}_{\text{CL}} = \mathcal{L}_{\text{node}} + \alpha \mathcal{L}_{\text{edge}} + \beta \mathcal{L}_{\text{cross}}, \tag{15}$$

where α and β balance the contributions of different structural granularities. This multi-level objective encourages consistency between corresponding representations while preserving structural distinctions across nodes, hyperedges, and views.

Higher-Order Hyperbolic Fusion and Downstream Classification Prediction

During the forward pass, the hyperbolic representations are passed to the higher-order propagation backbone. A feature adaptation layer first transforms the representations into the required latent space, after which hyperbolic self-attention is used to estimate pairwise interaction weights (Chen et al., 2022):

$$a_{\text{ij}} = \text{Softmax}_j \left(\frac{\langle \mathbf{z}_{v,i}, \mathbf{z}_{v,j} \rangle_{\mathcal{L}}}{\sqrt{d}} \right). \tag{16}$$

The resulting attention coefficients are used to aggregate higher-order information while remaining in the hyperbolic space. Specifically, the Einstein midpoint is employed to combine neighboring representations:

$$\mathbf{z}_{v,i}^{(\text{out})} = MidPoint\left(\{ a_{\text{ij}} \mathbf{z}_{v,j} \}_{j \in \mathcal{N}_i} \right), \tag{17}$$

where $\mathcal{N}_i$ denotes the relevant higher-order neighborhood of node v_i. This operation avoids repeatedly projecting representations into Euclidean space during aggregation and thereby preserves the underlying hyperbolic geometry.

Finally, the aggregated representations are mapped to the tangent space through $\log_o^k(\cdot)$ and fed into a Euclidean multilayer perceptron to obtain node-level predictions:

$$\widehat{\mathbf{Y}} = Softmax\left(\text{MLP}\left(\log_{\mathbf{o}}(\mathbf{Z}^{(\text{out})}) \right) \right). \tag{18}$$

CLhypformer is optimized jointly using the supervised node-classification objective and the topology-aware contrastive objective:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{NLL}}(\mathbf{Y}, \widehat{\mathbf{Y}}) + \lambda \mathcal{L}_{\text{CL}}, \tag{19}$$

where λ controls the contribution of the self-supervised objective.

3.2 Hyperbolic KNN-Based Undirected Hypergraph Construction

CLhypformer constructs the initial hypergraph directly in the Lorentz manifold, ensuring consistency between the geometry used for representation learning and the metric used for neighborhood formation.

Let the initial hyperbolic node representations be

$$X_{\mathcal{L}} = \{x_i\}_{i=1}^{N} \subset \mathcal{L}^d. \tag{20}$$

For each pair of nodes v_i and v_j, their hyperbolic geodesic distance is computed as

$$d_{\mathcal{L}}(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{\sqrt{\kappa}} \text{arcosh}\left(-\kappa \langle \mathbf{x}_i, \mathbf{x}_j \rangle_{\mathcal{L}} \right). \tag{21}$$

For each center node v_j, we identify its K nearest neighbors according to this hyperbolic distance:

$$\mathcal{N}_K(v_j) = \text{TopK}_{i \neq j}\left(-d_{\mathcal{L}}(\mathbf{x}_i, \mathbf{x}_j) \right) \tag{22}$$

Each center node then induces one hyperedge containing itself and its K nearest hyperbolic neighbors:

$$e_j = \{v_j\} \cup \mathcal{N}_K(v_j). \tag{23}$$

Thus, N center nodes induce $M = N$ hyperedges, yielding

$$\mathcal{E} = \{e_1, e_2, \ldots, e_N\}. \tag{24}$$

The corresponding incidence matrix $H \in \{0,1\}^{N \times N}$ is defined as

$$\mathbf{H}_{i,j} = \begin{cases} 1, & \mathrm{v}_i \in e_j \\ 0, & \text{otherwise} \end{cases} \tag{25}$$

Here, $H_{i,j} = 1$ indicates that node v_i belongs to the hyperedge centered at v_j. Since each hyperedge is defined as an unordered set of nodes, the resulting higher-order relation is inherently undirected. Consequently, no additional transposition-based symmetrization is required for the incidence matrix.

3.3 Controlled Degraded Variant and Joint Optimization

To isolate the contribution of topology-aware contrastive learning, we construct a controlled degraded variant, denoted by CLhypformer−. This variant removes the self-supervised components, including the topology-aware discriminator, asymmetric dual-view generation, and three-level contrastive objective, while retaining the Hyperbolic KNN-based hypergraph construction and the hyperbolic Transformer backbone. CLhypformer− is therefore trained exclusively with the supervised node-classification objective. For the full model, the supervised and self-supervised objectives are jointly optimized according to Eq. (19).

4 Results

4.1 Overall Performance Analysis

To evaluate the generalization capability of CLhypformer, we conduct experiments on six widely used graph benchmark datasets, including Cora and Citeseer. The results are summarized in Supplementary Table 1 and Figure 2. CLhypformer consistently achieves competitive or superior performance across the evaluated benchmarks, outperforming a range of representative baselines, including MLP, GCN, DHG-Bench, and Hypformer (Taud & Mas, 2018; Kipf & Welling, 2017; Li et al., 2026; Yang et al., 2024; Feng et al., 2024; Veličković et al., 2018). These results demonstrate the effectiveness of jointly modeling hyperbolic geometry and higher-order topology for graph representation learning.

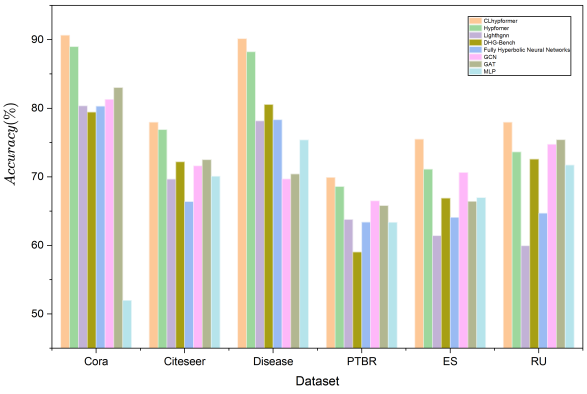


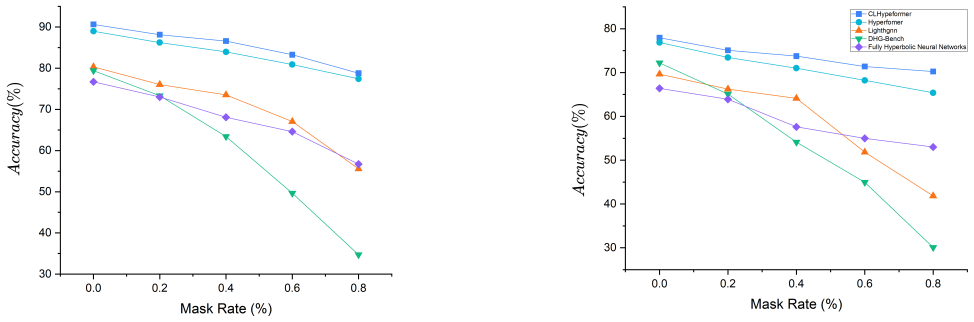
Figure 2: **Performance comparison of different models across benchmark datasets**. Node classification accuracy on six benchmark datasets; the legend identifies the evaluated models.

Hyperbolic representations provide a geometry better suited to hierarchical organization, whereas hyperbolic KNN-based hypergraph construction explicitly incorporates higher-order neighborhood relationships. In addition, topology-aware contrastive learning introduces structural consistency constraints across perturbed views, reducing the dependence of representation learning on potentially noisy or incomplete connectivity.

4.2 Robustness Analysis

Real-world graph data are often affected by missing edges, incomplete observations, or structural perturbations (Ju et al., 2025). To assess the robustness of CLhypformer to structural degradation, we progressively increase the structural masking rate and evaluate the resulting node classification performance. The results are reported in Supplementary Table 2 and Figure 3.

As the masking rate increases, the performance of all evaluated models gradually deteriorates, reflecting the increasing difficulty of recovering task-relevant information from incomplete topology. However, CLhypformer exhibits a comparatively slower performance degradation and maintains competitive performance under severe masking conditions. This result indicates that the proposed framework is less sensitive to structural corruption than the compared methods.



(a) Cora (b) Citeseer

Figure 3: **Robustness evaluation under varying levels of structural masking**. Classification accuracy of CLhypformer and baseline models as the structural masking rate increases on the evaluated benchmark datasets.

4.3 Ablation Study

To quantify the contribution of the major components of CLhypformer, we conduct controlled ablation experiments on the Cora dataset. In particular, we compare the complete model with the degraded variant, CLhypformer−, while separately examining the effects of hyperbolic KNN-based topology construction and topology-aware hyperbolic contrastive learning, as shown in Figure 4.

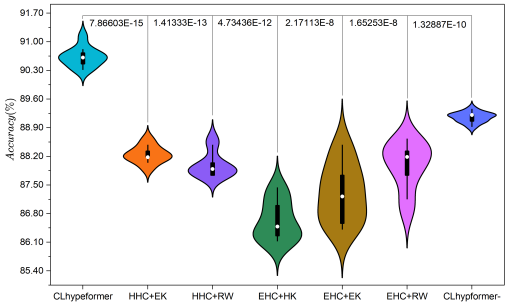
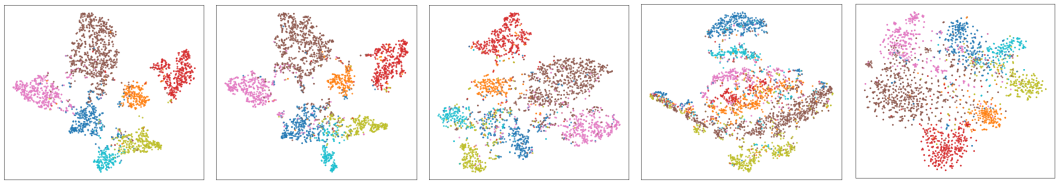


Figure 4: **Statistical significance analysis among model variants.** Pairwise statistical significance of the performance differences among the evaluated variants.

Supplementary Table 3 shows that hyperbolic KNN yields the highest accuracy when paired with hyperbolic contrastive learning. Under Euclidean contrastive learning, however, hyperbolic KNN performs worse than Euclidean KNN and random walks. These results support the importance of coupling topology construction with the contrastive representation space. Figure 4 visualizes the performance distributions and significance annotations of the compared variants.

More importantly, CLhypformer consistently outperforms CLhypformer−, despite both models sharing the same hyperbolic representation space and hypergraph Transformer backbone. The additional improvement demonstrates that topology-aware contrastive learning provides complementary information beyond supervised classification. In particular, the asymmetric topology-aware views and multi-level contrastive objective encourage representations to remain consistent across different structural perturbations, thereby improving the discriminative quality and structural stability of the learned embeddings.

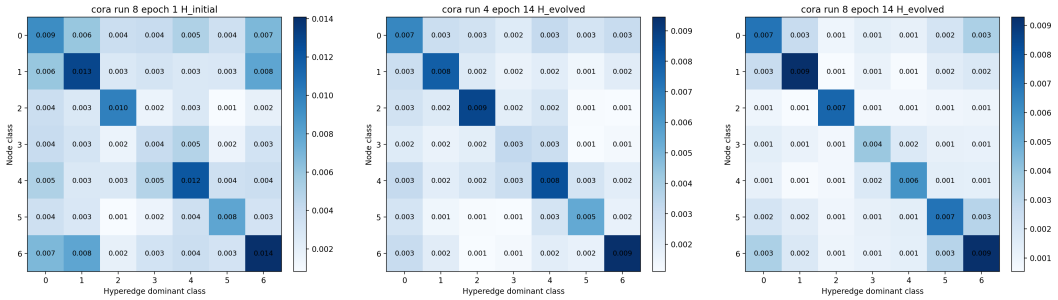
4.4 Qualitative Analysis and Topological Evolution



(a) CLhypformer (b) Hypformer (c) Lighthgnn (d) DHG-Bench (e) FHNN

Figure 5: Visualization of node embeddings learned by different models. The embeddings are projected into a two-dimensional space to compare their intra-class compactness and inter-class separability.

As shown in Figure 5, the node embeddings learned by the evaluated models differ in intra-class compactness and inter-class separability. CLhypformer forms more coherent class-specific groups with fewer ambiguous overlaps, indicating that the joint geometry–topology modeling improves the discriminative organization of the learned representation space.



(a) Initial (b) Hypformer (c) CLhypformer

Figure 6: Visualization of hypergraph topology through the incidence matrix. The incidence matrices illustrate the structural organization learned by different models, with CLhypformer (c) exhibiting a more distinct block structure and reduced off-diagonal connectivity.

Figure 6 compares the initial incidence matrix with those learned by Hypformer and CLhypformer. The initial topology contains relatively dispersed connections, including structurally ambiguous relationships across different node groups. Compared with Hypformer, CLhypformer produces a more distinct block-like organization and fewer cross-group associations at the selected training stage, suggesting that the proposed topology-aware learning objective improves the consistency between learned representations and task-relevant structural patterns.

5 Conclusion

We presented CLhypformer, a topology-aware hyperbolic hypergraph Transformer for graph representation learning. CLhypformer integrates hyperbolic KNN-based hypergraph construction with topology-aware hyperbolic contrastive learning and higher-order hypergraph Transformer propagation within a unified framework.

Extensive experiments demonstrate that CLhypformer achieves competitive or superior performance across multiple node classification benchmarks and exhibits improved robustness under severe structural masking. Ablation and qualitative analyses further show that both hyperbolic topology construction and topology-aware contrastive alignment contribute to the overall performance.

AI USE STATEMENT

Generative AI tools were used solely for language polishing, LaTeX formatting, and assistance with code debugging. They were not used for research ideation, experimental design or execution, data generation or analysis, or figure generation.

REFERENCES

Alessia Antelmi, Gennaro Cordasco, Mirko Polato, Vittorio Scarano, Carmine Spagnuolo, and Dingqi Yang. A survey on hypergraph representation learning. ACM Computing Surveys, 56 (1): 24:1–24:38, 2023.

Ines Chami, Rex Ying, Christopher Ré, and Jure Leskovec. Hyperbolic graph convolutional neural networks. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Tianyi Chen, Haobo Wang, Kai Tang, Gengyu Lyu, Tianlei Hu, Gang Chen, Hong Ma, and Meixiang Xiang. Hyperbolic enhanced representation learning for incomplete multi-view clustering. arXiv preprint arXiv:2604.16959, 2026.

Weize Chen, Xu Han, Yankai Lin, Hexu Zhao, Zhiyuan Liu, Peng Li, Maosong Sun, and Jie Zhou. Fully hyperbolic neural networks. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, pp. 5672–5686, 2022.

Ching-Yao Chuang, Joshua W. Robinson, Yen-Chen Lin, Antonio Torralba, and Stefanie Jegelka. Debiased contrastive learning. In *Advances in Neural Information Processing Systems*, volume 33, 2020.

Yifan Feng, Haoxuan You, Zizhao Zhang, Rongrong Ji, and Yue Gao. Hypergraph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 3558–3565, 2019.

Yifan Feng, Yihe Luo, Shihui Ying, and Yue Gao. LightHGNN: Distilling hypergraph neural networks into MLPs for 100× faster inference. In *The Twelfth International Conference on Learning Representations*, 2024.

Xingcheng Fu, Jianxin Li, Jia Wu, Qingyun Sun, Cheng Ji, Senzhang Wang, Jiajun Tan, Hao Peng, and Philip S Yu. Ace-hgnn: Adaptive curvature exploration hyperbolic graph neural network. In 2021 IEEE international conference on data mining (ICDM), pp. 111–120. IEEE, 2021.

Wei Ju, Siyu Yi, Yifan Wang, Zhiping Xiao, Zhengyang Mao, Hourun Li, Yiyang Gu, Yifang Qin, Nan Yin, Senzhang Wang, et al. A survey of graph neural networks in real world: Imbalance, noise, privacy and ood challenges. IEEE transactions on pattern analysis and machine intelligence, 2025.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*, 2017.

Marc Law, Renjie Liao, Jake Snell, and Richard Zemel. Lorentzian distance learning for hyperbolic representations. In International conference on machine learning, pp. 3672–3681. PMLR, 2019.

Dongjin Lee and Kijung Shin. I'm me, we're us, and I'm us: Tri-directional contrastive learning on hypergraphs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 8456–8464, 2023.

Fan Li, Xiaoyang Wang, Wenjie Zhang, Ying Zhang, and Xuemin Lin. DHG-Bench: A comprehensive benchmark for deep hypergraph learning. In *International Conference on Learning Representations*, 2026.

Zhuo Li, Yuena Lin, Yipeng Wang, Wenmao Liu, Mingliang Yu, Zhen Yang, and Gengyu Lyu. Critical node-aware augmentation for hypergraph contrastive learning. In IJCAI, pp. 5671–5679, 2025.

Maximilian Nickel and Douwe Kiela. Learning continuous hierarchies in the Lorentz model of hyperbolic geometry. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of Proceedings of Machine Learning Research, pp. 3779–3788. PMLR, 2018.

Wei Peng, Tuomas Varanka, Abdelrahman Mostafa, Henglin Shi, and Guoying Zhao. Hyperbolic deep neural networks: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44 (12): 10023–10044, 2022.

Hind Taud and Jean-François Mas. Multilayer perceptron (MLP). In Geomatic Approaches for Modeling Land Change Scenarios, pp. 451–455. Springer, 2018.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*, 2018.

Menglin Yang, Min Zhou, Tong Zhang, Jiahong Liu, Zhihao Li, Lujia Pan, Hui Xiong, and Irwin King. Hyperbolic graph neural networks: A review of methods and applications. arXiv preprint arXiv:2202.13852, 2022.

Menglin Yang, Harshit Verma, Delvin Ce Zhang, Jiahong Liu, Irwin King, and Rex Ying. Hypformer: Exploring efficient transformer fully in hyperbolic space. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 3770–3781, 2024.

Hai-Cheng Yi, Zhu-Hong You, De-Shuang Huang, and Chee Keong Kwoh. Graph representation learning in bioinformatics: trends, methods and applications. Briefings in Bioinformatics, 23 (1): bbab340, 2022.