BrainHTF: Learning Causal Graph Representation of Brain Connectome with Hyperbolic Transformer

Qiyu Sun[1], Zhenlin Mao[1], Ruiqing Feng[1], Jiashuang Huang[2], and Mingliang Wang[1,⋆]

[1] School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China
[2] School of Artificial Intelligence and Computer Science, Nantong University, Nantong 226019, China
wml489@nuist.edu.cn

Abstract. Brain connectome provides a comprehensive map of neural connections within the brain, which is fundamental to understanding brain function and dysfunction. Recently, transformer-based methods have significantly advanced the analysis of brain connectome. However, these methods mainly rely on learning node-level representations, overlooking the critical role of brain connectivity (edge) to elucidate the mechanism of brain disorders, which limits their interpretability. Besides, the brain connectome exhibits a modular and hierarchical topology, while traditional Euclidean-based methods struggle to model these hierarchical structures without high distortion, limiting their representational capacity. To address these limitations, we propose BrainHTF, a novel hyperbolic manifold-based Transformer framework tailored for learning causal graph representation of brain connectome. Specifically, we first introduce a mask generator to disentangle the brain connectome into causal and bias subgraphs. Then, a hyperbolic transformer is utilized to learn their respective disentangled representations in hyperbolic space. Finally, causal intervention is conducted on representations for discovering invariant causal patterns for prediction. Experimental results on the ABIDE and ADNI datasets demonstrate that our BrainHTF consistently outperforms several state-of-the-art methods in brain disease detection while also identifying interpretable biomarkers. The code is available at https://github.com/MYzzZ123/BrainHTF.

Keywords: Brain Connectome · Hyperbolic Transformer · Brain Disease · Representation Learning.

1 Introduction

Brain connectome has intrigued neuroscientists for its potential to elucidate intricate organizations of human brain and underlying mechanisms of neurological

⋆ Corresponding author

disorders [9,12]. At the macro-scale, the brain connectome is modeled as a brain network or graph, where nodes represent regions-of-interest (ROIs) and edges represent connections between them. Functional magnetic resonance imaging (fMRI), a noninvasive neuroimaging modality, measures neural activity through blood-oxygen-level-dependent (BOLD) signals. Functional connectivity (FC) is derived by quantifying statistical dependencies between pairs of BOLD signals. Numerous studies have demonstrated that individuals with autism spectrum disorder (ASD) and Alzheimer's disease (AD) exhibit abnormal FC patterns [5,15]. FC-based analysis of brain connectome has been one of the key approaches to decipher underlying mechanisms of brain diseases.

Nowadays, Transformer-based models have led to significant advances in learning representation of the brain connectome. For instance, Brain Network Transformer (BrainNetTF) [10] utilizes a transformer encoder with an orthonormal clustering readout layer to learn graph representation of connectome for brain disease detection. Long-range Brain Graph Transformer (ALTER) [17] captures long-range dependencies between ROI pairs via biased random walk and adaptively integrates them with short-range dependencies, facilitating an integrated understanding of communication across the whole brain. However, existing Transformer-based methods primarily focus on learning node-level representations, neglecting the critical contribution of brain connectivity (edge) to understanding the mechanisms underlying brain disorders [7]. Besides, most existing methods are confined to learn representations within Euclidean geometry, which struggles to model the modular and hierarchical network structure within brain connectome due to inherent spatial constraints, resulting in limited representational capacity.

To address these issues, we propose BrainHTF, a novel hyperbolic manifold-based Transformer framework tailored for learning causal graph representation of brain connectome. Specifically, given a brain graph, we propose to split it into causal and bias subgraphs via a subgraph generator. Then, a hyperbolic transformer is employed to effectively learn their respective disentangled graph representations based on hyperbolic geometry, which models the hierarchical structure of brain connectome with high fidelity. Finally, causal intervention is conducted on representations to discover invariant causal patterns for prediction. We conduct extensive experiments on two public fMRI datasets (i.e., ABIDE and ADNI), with experimental results demonstrating that BrainHTF not only outperforms state-of-the-art methods in brain disease classification tasks, but also hold promise in identifying interpretable biomarkers. The major contributions of this work are summarized as follows:

1) We propose a novel framework, called BrainHTF, which learns causal graph representation of brain connectome based on hyperbolic geometry for brain disease detection.
2) We design a hyperbolic transformer that embeds the brain's hierarchical topology into Poincaré ball hyperbolid to mitigate Euclidean representational limitations. To the best of our knowledge, this is the first work to introduce a hyperbolic transformer for brain connectome analysis.

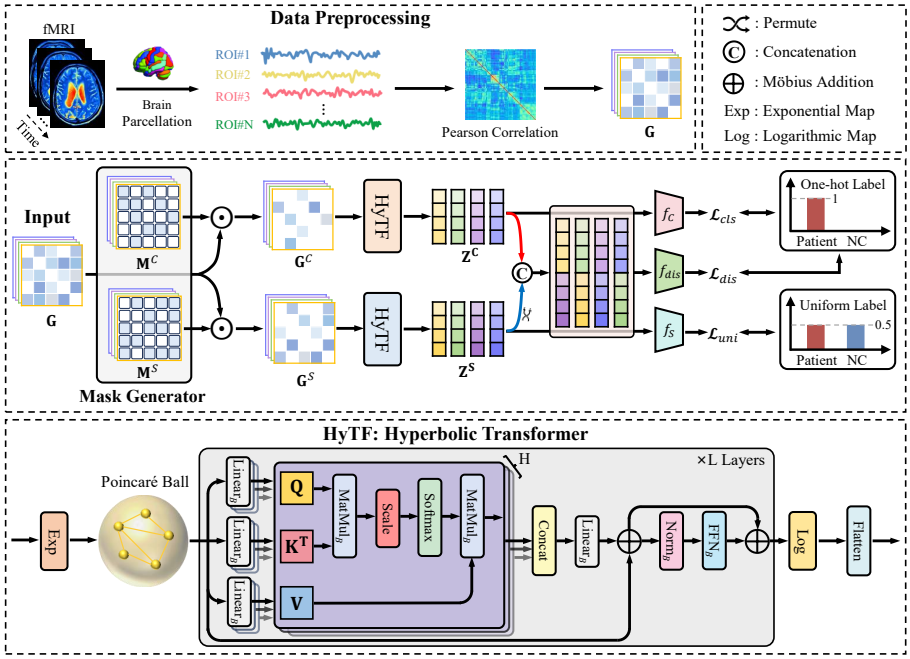


Fig. 1. The overall framework of the proposed BrainHTF.

3) Extensive experiments on two widely-used fMRI datasets demonstrate that our BrainHTF outperforms all baseline methods for brain disease detection, while holds potential in identifying interpretable biomarkers.

2 Preliminaries

Notations. In this paper, $\mathcal{R}^n$ and $\mathcal{B}_c^n$ represents the n-dimensional Euclidean space and the n-dimensional Poincaré ball model with negative curvature $-c$, respectively. $\oplus$ denotes the Möbius addition, and $\log(\cdot)$ and $\exp(\cdot)$ denote the logarithmic and exponential maps, respectively. $\mathcal{G} = \{\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_I\} \in \mathcal{R}^{I \times N \times N}$ denotes a batch of brain graphs, where I and N denotes the number of subjects and ROIs, respectively.

Hyperbolic Geometry. Hyperbolic geometry is a smooth Riemannian manifold with constant negative curvature [16]. Several isometric models have been proposed to represent hyperbolic space [1]. In this paper, we select the Poincaré ball model as the geometric foundation of this study. The Poincaré ball model $(\mathcal{B}_c^n, g_c^n)$ of an n-dimensional hyperbolic space is defined as:

$$\mathcal{B}_c^n = \{\mathbf{x} \in \mathcal{R}^n : c\|\mathbf{x}\|^2 < 1, c > 0\}, \tag{1}$$

with constant negative curvature $-c$ and the hyperbolic metric g_c^n.

Hyperbolic Linear Transformation. Let $\mathbf{W}$ be a weight matrix in the tangent space, we can define hyperbolic linear transformation as:

$$\text{Linear}_{\mathcal{B}}(\mathbf{x}) = (\mathbf{W} \otimes_c \mathbf{x}) \oplus_c \mathbf{b}, \tag{2}$$

where $\mathbf{W} \otimes_c \mathbf{x} = \exp_{\mathbf{0}}^c(\mathbf{W} \log_{\mathbf{0}}^c(\mathbf{x}))$, $\mathbf{b}$ is the bias term in the hyperbolic space.
Hyperbolic Layer Normalization. Let $\text{LN}_{\mathcal{E}}(\cdot)$ be the layer norm in the Euclidean space, we can define hyperbolic layer normalization as:

$$\text{LN}_{\mathcal{B}}(\mathbf{x}) = \exp_{\mathbf{0}}^c\Big(\text{LN}_{\mathcal{E}}(\log_{\mathbf{0}}^c(\mathbf{x}))\Big). \tag{3}$$

3 Method

3.1 Causal Subgraph Generator

Given a batch of brain graphs $\mathcal{G} = \{\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_I\} \in \mathcal{R}^{I \times N \times N}$, we design a mask generator to split each brain graph into causal and bias subgraphs. Specifically, given a brain graph $\mathbf{G}$, where $\mathbf{g}_i$ and $\mathbf{g}_j$ represent the feature of node i and j, we first use a Multi-layer Perceptron (MLP) followed by a sigmoid function to obtain the edge importance matrix $\mathbf{D} \in \mathcal{R}^{N \times N}$, where $\mathbf{D}_{ij}$ indicates the importance of edge (i, j):

$$\mathbf{D}_{ij} = \sigma(\text{MLP}(\mathbf{g}_i \,\|\, \mathbf{g}_j)), \tag{4}$$

where $\|$ is the concatenation operator, and $\sigma(\cdot)$ is the sigmoid function. Then, we select the edges with highest importance to construct the binarized causal mask $\mathbf{M}^C \in \{0,1\}^{N \times N}$, while the remaining edges to derive the binarized bias mask $\mathbf{M}^S \in \{0,1\}^{N \times N}$:

$$\mathbf{M}^C = \mathbb{I}\big(\text{Top}_{p_c}(\mathbf{D})\big), \quad \mathbf{M}^S = \mathbf{1} - \mathbf{M}^C, \tag{5}$$

where $\mathbf{1} \in \mathcal{R}^{N \times N}$ denotes the all-one matrix. The indicator function $\mathbb{I}(\cdot)$ returns 1 if the condition is satisfied, 0 otherwise. $\text{Top}_{p_c}(\cdot)$ selects the top-K edges with $K = p_c \times N \times N$, and $p_c\%$ is a hyperparameter (e.g., 40%). With these binarized masks, we can filter $\mathbf{G}$ as below:

$$\mathbf{G}^C = \mathbf{G} \odot \mathbf{M}^C, \quad \mathbf{G}^S = \mathbf{G} \odot \mathbf{M}^S, \tag{6}$$

where $\odot$ denotes the element-wise product. $\mathbf{G}^C$ and $\mathbf{G}^S \in \mathcal{R}^{N \times N}$ represent the causal and bias subgraph features, respectively.

3.2 Learning Disentangled Graph Representations

To model the intrinsic hierarchical structure of the brain connectome, we propose a hyperbolic Transformer that learns disentangled graph representations of $\mathbf{G}^C$

and $\mathbf{G}^S$ based on hyperbolic geometry. Concretely, given $\mathbf{G}$, the full process of the l-th layer in our transformer encoder is formulated as:

$$\mathbf{Q}^{l,h}, \mathbf{K}^{l,h}, \mathbf{V}^{l,h} = \mathbf{H}^{l-1,h} \otimes_c \left(\mathbf{W}_Q^{l,h}, \mathbf{W}_K^{l,h}, \mathbf{W}_V^{l,h} \right), \tag{7}$$

$$\mathbf{S}^{l,h} = \text{softmax} \left(\frac{\log_{\mathbf{0}}^c(\mathbf{Q}^{l,h}) \log_{\mathbf{0}}^c(\mathbf{K}^{l,h})^\top}{\sqrt{d_K^{l,h}}} \right), \tag{8}$$

$$\mathbf{O}^{l,h} = \exp_{\mathbf{0}}^c(\mathbf{S}^{l,h} \log_{\mathbf{0}}^c(\mathbf{V}^{l,h})), \quad \mathbf{O}^l = \left(\|_{h=1}^H \mathbf{O}^{l,h} \right) \otimes_c \mathbf{W}_O^l, \tag{9}$$

$$\mathbf{H}' = \text{LN}_{\mathcal{B}}(\mathbf{O}^l \oplus_c \mathbf{H}^{l-1}), \quad \mathbf{H}^l = \text{FFN}_{\mathcal{B}}(\mathbf{H}') \oplus_c \mathbf{H}', \tag{10}$$

where $\mathbf{O}^0 = \exp_{\mathbf{0}}^c(\mathbf{G})$, H is the number of heads, l is the layer index, $\mathbf{S}^{l,h}$ is the attention score, and $\mathbf{W}_Q^{l,h}$, $\mathbf{W}_K^{l,h}$, $\mathbf{W}_V^{l,h}$, and $\mathbf{W}_O^l$ are learnable manifold parameter matrices. $d_K^{l,h}$ is the first dimension of $\mathbf{W}_K^{l,h}$. $\text{FFN}_{\mathcal{B}}(\cdot)$ denotes the feed-forward network, consisting of two hyperbolic linear layers. Finally, the output of hyperbolic transformer encoder is projected to the tangent space, followed by flattening into a vector as the final representation.

Following the above encoding process, we learn disentangled representations of the causal and bias subgraphs and make predictions via classifiers:

$$\mathbf{Z}^C = \Phi_C(\mathbf{G}^C), \quad \mathbf{y}_{\mathbf{G}}^C = f_C(\mathbf{Z}^C), \quad \mathbf{Z}^S = \Phi_S(\mathbf{G}^S), \quad \mathbf{y}_{\mathbf{G}}^S = f_S(\mathbf{Z}^S), \tag{11}$$

where $\mathbf{Z}^C$ and $\mathbf{Z}^S$ are the representations of the causal and bias subgraphs. $\Phi_C(\cdot)$ and $\Phi_S(\cdot)$ are the causal and bias transformer encoders. $\mathbf{y}_{\mathbf{G}}^C$ and $\mathbf{y}_{\mathbf{G}}^S$ are the predictions of the causal classifier $f_C(\cdot)$ and the bias classifier $f_S(\cdot)$, respectively.

3.3 Loss Function

The causal subgraph is derived to capture causal representation, so we assign its representation to the ground-truth label $\mathbf{y}_{\mathbf{G}}$. Thus, we define the classification loss as:

$$\mathcal{L}_{cls} = -\frac{1}{|\mathcal{G}|} \sum_{\mathbf{G} \in \mathcal{G}} \mathbf{y}_{\mathbf{G}}^\top \log \left(\mathbf{y}_{\mathbf{G}}^C \right). \tag{12}$$

The bias subgraph aims to approach the spurious patterns that do not contribute to the classification. Hence, we enforce its prediction to follow a uniform distribution over all categories and define the uniform classification loss as:

$$\mathcal{L}_{uni} = \frac{1}{|\mathcal{G}|} \sum_{\mathbf{G} \in \mathcal{G}} \text{KL}\left(\mathbf{y}_{uni}, \mathbf{y}_{\mathbf{G}}^S \right), \tag{13}$$

where $\text{KL}(\cdot)$ is the KL divergence, and $\mathbf{y}_{\text{unif}}$ denotes the uniform distribution.

To capture invariant causal patterns from brain graphs, we conduct interventions on bias representations. Specifically, we first randomly permute $\mathbf{Z}^S$ in a

batch and obtain $\mathbf{Z}^{dis} = \mathbf{Z}^C \| \mathbf{Z}^{S'}$, where $\mathbf{Z}^{S'}$ represents randomly permuted bias vectors of $\mathbf{Z}^S$. With the generated interventions, we define the following loss:

$$\mathbf{y}_{\mathbf{G}}^{dis} = f_{dis}(\mathbf{Z}^{dis}), \tag{14}$$

$$\mathcal{L}_{dis} = -\frac{1}{|\mathcal{G}|} \sum_{\mathbf{G} \in \mathcal{G}} \mathbf{y}_{\mathbf{G}}^{\top} \log \left(\mathbf{y}_{\mathbf{G}}^{dis} \right), \tag{15}$$

where $\mathbf{y}_{\mathbf{G}}^{dis}$ is the prediction produced by the classifier $f_{dis}(\cdot)$.

Finally, the objective function is defined as the weighted sum of the losses:

$$\mathcal{L} = \mathcal{L}_{cls} + \lambda_1 \mathcal{L}_{uni} + \lambda_2 \mathcal{L}_{dis}, \tag{16}$$

where λ_1 and λ_2 are parameters to regulate the contribution of disentanglement and causal intervention, respectively.

4 Experiments

4.1 Experimental Settings

Datasets and Preprocessing. We evaluate our method on two public fMRI datasets: (1) *Autism Brain Imaging Data Exchange* (ABIDE) [4], which contains 571 ASD patients and 531 NCs. (2) *Alzheimer's Disease Neuroimaging Initiative* (ADNI) [8], which contains 99 AD patients and 154 NCs. Both datasets were preprocessed following the standard pipeline. Incomplete scan images were excluded from our study to ensure data quality.

Baselines. To evaluate the effectiveness of our method, we compare BrainHTF with three types of baselines: (1) GNN-based methods: BrainUSL [18] and BrainIB [19]. (2) Transformer-based methods: BrainNetTF [10] and ALTER [17]. (3) Hyperbolic manifold-based methods: HGCN [2] and H2H-GCN [3].

Metrics. We adopt a five-fold cross-validation to assess the performance of different methods. Four metrics are used for performance evaluation: accuracy (ACC), sensitivity (SEN), specificity (SPE), and area under the receiver operating characteristic curve (AUC), respectively.

Implementation Details. BrianHTF is implemented based on Pytorch and trained on an NVIDIA GeForce RTX 4090 GPU. In the hyperbolic transformer encoder, the number of attention heads and layers are both set to 2, and the curvature is set to 3. All classifier are based on MLP. The parameters λ_1 and λ_2 in Eq. (16) are both set to 0.1. During the training process, the RiemannianAdam optimizer is used with a learning rate of 1e-5 and a weight decay of 1e-5. The batch size is 8 and the epoch is 100.

4.2 Results Analysis

Table 1 reports the results of all methods in ASD vs. NC and AD vs. NC classification tasks. Several key observations can be drawn from this table. *First*,

Transformer-based methods consistently outperforms GNN-based methods (i.e., BrainUSL and BrainIB) on ABIDE and ADNI datasets. For example, BrainNetTF achieves the ACC values of 64.69% and 85.74% in ASD vs. NC and AD vs. NC classification tasks, which is significantly higher than 61.24% and 81.84% yielded by BrainIB. *Second*, hyperbolic manifold-based methods generally achieve performance comparable to GNN-based methods, but remain inferior to Transformer-based methods. This is likely because these manifold-based methods are not specifically designed for brain connectome analysis, making it difficult to learn rich representations for prediction. *Finally*, our BrainHTF achieves the optimal performance across both datasets. In particular, BrainHTF reaches the average ACC and AUC of 80.72% and 84.46%, yielding gains of 3.26% and 1.63%, respectively, over the suboptimal ALTER at 77.46% and 82.83%. These notable improvements can be attributed two key contributions: 1) Tailored mask generator equipped with causal intervention can effectively extract causal connectivity patterns associated with ASD/AD from brain connectome; 2) Modeling hierarchical topology of brain connectome in hyperbolic space enables learning rich representations for prediction.

Table 1. Experimental results of different methods on ABIDE and ADNI datasets.

Dataset	Method	ACC(%) ↑	SEN(%) ↑	SPE(%) ↑	AUC(%) ↑
ABIDE	BrainUSL [18]	64.10±4.69	72.89±8.72	54.46±7.98	69.10±5.29
	BrainIB [19]	61.24±3.72	66.29±9.31	55.69±3.77	64.59±3.65
	BrainNetTF [10]	64.69±2.78	65.90±7.37	63.35±6.48	70.34±3.92
	ALTER [17]	67.16±2.72	70.81±4.28	63.15±1.80	71.81±3.26
	HGCN [2]	62.82±2.88	68.91±12.64	56.14±12.71	68.18±3.05
	H2H-GCN [3]	64.10±2.78	66.44±10.70	61.46±12.68	70.59±4.61
	BrainHTF	**71.30±2.76**	**74.58±4.88**	**67.72±4.73**	**74.03±3.96**
ADNI	BrainUSL [18]	81.03±5.09	74.74±5.50	85.12±5.92	89.43±3.52
	BrainIB [19]	81.84±3.76	76.89±11.97	85.12±8.52	87.26±5.60
	BrainNetTF [10]	85.74±3.53	81.79±6.86	88.34±5.20	92.17±3.65
	ALTER [17]	87.76±5.48	82.89±10.23	90.95±6.24	93.84±2.48
	HGCN [2]	79.44±2.75	64.58±9.22	88.99±5.59	84.82±4.22
	H2H-GCN [3]	79.84±7.37	78.89±11.48	80.47±7.51	88.14±5.47
	BrainHTF	**90.13±2.15**	**83.89±3.27**	**94.19±3.76**	**94.89±2.42**

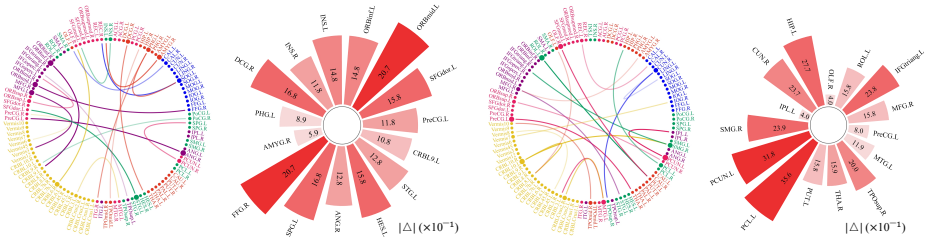
4.3 Ablation Study

To evaluate the effectiveness of key components, we compare our BrainHTF with their variants: (1) "w/o $H_{\mathcal{B}}$" that replaces our hyperbolic transformer with a standard transformer, (2) "w/o $M_{\mathcal{D}}$" that removes the causal mask generator for disentanglement, and (3) "w/o $C_{\mathcal{I}}$" that removes the causal intervention.

As reported in Table 2, "w/o $H_{\mathcal{B}}$" shows a significant performance drop, showing the potential of hyperbolic geometry in modeling hierarchical brain con-

Table 2. Ablation studies of different components on ABIDE and ADNI datasets.

Dataset	Method	ACC(%) ↑	SEN(%) ↑	SPE(%) ↑	AUC(%) ↑
ABIDE	w/o $H_{\mathcal{B}}$	66.17±2.20	69.68±4.08	62.33±4.05	71.38±3.16
	w/o $M_{\mathcal{D}}$	65.58±2.55	69.30±5.13	61.49±1.77	71.24±3.61
	w/o $C_{\mathcal{I}}$	65.28±2.03	68.16±4.33	62.12±2.98	71.17±2.09
	BrainHTF	**71.30±2.76**	**74.58±4.88**	**67.72±4.73**	**74.03±3.96**
ADNI	w/o $H_{\mathcal{B}}$	87.76±3.77	81.89±8.02	91.59±3.26	93.93±2.22
	w/o $M_{\mathcal{D}}$	88.93±2.93	**85.79±6.06**	90.95±4.71	**95.13±2.31**
	w/o $C_{\mathcal{I}}$	87.36±1.93	82.89±4.96	90.30±4.05	94.45±3.03
	BrainHTF	**90.13±2.15**	83.89±3.27	**94.19±3.76**	94.89±2.42



(a) ASD vs. NC (b) AD vs. NC

Fig. 2. Visualization of the causal connectivity patterns and the perturbation analysis.

nectome. "w/o $M_{\mathcal{D}}$" consistently degrades the performance. This demonstrates the effectiveness of disentangling causal patterns from brain network for prediction. "w/o $C_{\mathcal{I}}$" shows the largest performance degradation on both datasets, which implies that causal intervention is beneficial for discovering causal patterns for brain disease detection. Notably, our BrainHTF achieves the overall best performance compare with their variants, proving that discovering causal patterns and modeling them in hyperbolic space is helpful for improving model's diagnostic performance.

4.4 Visualization Analysis

We investigated the identified causal connectivity patterns associated with ASD and AD classification. Specifically, we calculated the average values of causal connectivity in the edge importance matrix $\mathbf{D}$ in Eq. (4) across all folds, which served as the contribution of causal connectivities to ASD and AD identification. Besides, we conduct perturbations on selected ROIs among causal connectivities to investigate the association between the ROIs and disease. These results were presented in Fig. 2. From Fig. 2 (a), we can observe that the detected regions such as *left middle orbital frontal gyrus* (ORBmid.L), *left dorsolateral superior frontal gyrus* (SFGdor.L), and *left insula* (INS.L) play crucial role in ASD identification, consistent with previous studies [6,11,14]. As shown in Fig. 2 (b), *left hippocampus* (HIP.L) was identified, which was recognized as one of the most

affected areas in AD identification [13]. As illustrated in Fig. 2 (a) and (b) (right part), perturbations on causal connectivity lead to notable changes in prediction (e.g., $|\Delta| \approx 20.7 \times 10^{-1}$ and $|\Delta| \approx 27.7 \times 10^{-1}$ for ORBmid.L and HIP.L in ASD and AD classification, respectively), validating that our BrainHTF is capable of identifying key biomarkers for brain disease detection.

5 Conclusion

In this study, we propose BrainHTF, a novel Hyperbolic manifold-based Transformer framework tailored for learning causal graph representation of brain connectome. In BrainHTF, given a brain graph, we first utilize a mask generator to disentangle it into causal and bias subgraphs. Then, they are mapped to hyperbolic space for disentangled representations learning with hyperbolic transformers. Finally, causal intervention is conducted on bias representations to discover invariant causal patterns for prediction. Experimental results on the ABIDE and ADNI datasets demonstrate the superiority of BrainHTF in brain disorders detection and its ability of identifying interpretable biomarkers.

Acknowledgments. This work was supported by the National Natural Science Foundation of China under Grant Nos. 62472228 and 62471259.

Disclosure of Interests. The authors have no competing interests to declare relevant to this article's content.

References

1. Cannon, J.W., Floyd, W.J., Kenyon, R., Parry, W.R., et al.: Hyperbolic Geometry. Flavors of Geometry **31**(59-115), 2 (1997)
2. Chami, I., Ying, Z., Ré, C., Leskovec, J.: Hyperbolic Graph Convolutional Neural Networks. Advances in Neural Information Processing Systems **32** (2019)
3. Dai, J., Wu, Y., Gao, Z., Jia, Y.: A Hyperbolic-to-Hyperbolic Graph Convolutional Network. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 154–163 (2021)
4. Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., et al.: The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular Psychiatry **19**(6), 659–667 (2014)
5. Gu, Y., Lin, Y., Huang, L., Ma, J., Zhang, J., Xiao, Y., Dai, Z., Initiative, A.D.N.: Abnormal dynamic functional connectivity in alzheimer's disease. CNS Neuroscience & Therapeutics **26**(9), 962–971 (2020)
6. Habata, K., Cheong, Y., Kamiya, T., Shiotsu, D., Omori, I.M., Okazawa, H., Jung, M., Kosaka, H.: Relationship between sensory characteristics and cortical thickness/volume in autism spectrum disorders. Translational Psychiatry **11**(1), 616 (2021)
7. Insel, T.R., Cuthbert, B.N.: Brain disorders? Precisely. Science **348**(6234), 499–500 (2015)

8. Jack Jr, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., L. Whitwell, J., Ward, C., et al.: The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine **27**(4), 685–691 (2008)
9. Jiang, X., Shen, Y., Yao, J., Zhang, L., Xu, L., Feng, R., Cai, L., Liu, J., Chen, W., Wang, J.: Connectome analysis of functional and structural hemispheric brain networks in major depressive disorder. Translational Psychiatry **9**(1), 136 (2019)
10. Kan, X., Dai, W., Cui, H., Zhang, Z., Guo, Y., Yang, C.: Brain Network Transformer. Advances in Neural Information Processing Systems **35**, 25586–25599 (2022)
11. Kurochkin, I., Khrameeva, E., Tkachev, A., Stepanova, V., Vanyushkina, A., Stekolshchikova, E., Li, Q., Zubkov, D., Shichkova, P., Halene, T., et al.: Metabolome signature of autism in the human prefrontal cortex. Communications Biology **2**(1), 234 (2019)
12. de Lange, S.C., Scholtens, L.H., Initiative, A.D.N., van den Berg, L.H., Boks, M.P., Bozzali, M., Cahn, W., Dannlowski, U., Durston, S., Geuze, E., et al.: Shared vulnerability for connectome alterations across psychiatric and neurological brain disorders. Nature Human Behaviour **3**(9), 988–998 (2019)
13. Moreno-Jiménez, E.P., Flor-García, M., Terreros-Roncal, J., Rábano, A., Cafini, F., Pallas-Bazarra, N., Ávila, J., Llorens-Martín, M.: Adult hippocampal neurogenesis is abundant in neurologically healthy subjects and drops sharply in patients with alzheimer's disease. Nature Medicine **25**(4), 554–560 (2019)
14. Uddin, L.Q., Menon, V.: The anterior insula in autism: under-connected and under-examined. Neuroscience & Biobehavioral Reviews **33**(8), 1198–1203 (2009)
15. Yang, C., Wang, X.K., Ma, S.Z., Lee, N.Y., Zhang, Q.R., Dong, W.Q., Zang, Y.F., Yuan, L.X.: Abnormal functional connectivity of the reward network is associated with social communication impairments in autism spectrum disorder: A large-scale multi-site resting-state fMRI study. Journal of Affective Disorders **347**, 608–618 (2024)
16. Yang, M., Sun, Y., Wang, Q., Bozoki, A., Kohi, M., Liu, M.: Hyperbolic Kernel Graph Neural Networks for Neurocognitive Decline Analysis from Multimodal Brain Imaging. IEEE Transactions on Pattern Analysis and Machine Intelligence (2025)
17. Yu, S., Jin, S., Li, M., Sarwar, T., Xia, F.: Long-range Brain Graph Transformer. Advances in Neural Information Processing Systems **37**, 24472–24495 (2024)
18. Zhang, P., Wen, G., Cao, P., Yang, J., Zhang, J., Zhang, X., Zhu, X., Zaiane, O.R., Wang, F.: BrainUSL: Unsupervised Graph Structure Learning for Functional Brain Network Analysis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 205–214. Springer (2023)
19. Zheng, K., Yu, S., Li, B., Jenssen, R., Chen, B.: BrainIB: Interpretable brain network-based psychiatric diagnosis with graph information bottleneck. IEEE Transactions on Neural Networks and Learning Systems **36**(7), 13066–13079 (2025)