# Path-Based Heterogeneous Brain Transformer Network for Resting-State Functional Connectivity Analysis

Ruiyan Fang[1], Yu Li[1], Xin Zhang[1,2(✉)], Shengxian Chen[1], Jiale Cheng[1,3], Xiangmin Xu[1,2], Jieling Wu[4], Weili Lin[3], Li Wang[3], Zhengwang Wu[3], and Gang Li[3(✉)]

[1] School of Electronic and Information Engineering, South China University of Technology, Guangzhou, Guangdong, China
eexinzhang@scut.edu.cn
[2] Pazhou Lab, Guangzhou, China
[3] Department of Radiology and Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
gang_li@med.unc.edu
[4] Guangdong Maternal and Child Health Center, Guangzhou, China

**Abstract.** Brain functional connectivity analysis is important for understanding brain development, aging, sexual distinction and brain disorders. Existing methods typically adopt the resting-state functional connectivity (rs-FC) measured by functional MRI as an effective tool, while they either neglect the importance of information exchange between different brain regions or the heterogeneity of brain activities. To address these issues, we propose a Path-based Heterogeneous Brain Transformer Network (PH-BTN) for analyzing rs-FC. Specifically, to integrate the path importance and heterogeneity of rs-FC for a comprehensive description of the brain, we first construct the brain functional network as a path-based heterogeneous graph using prior knowledge and gain initial edge features from rs-FC. Then, considering the constraints of graph convolution in aggregating long-distance and global information, we design a Heterogeneous Path Graph Transformer Convolution (HP-GTC) module to extract edge features by aggregating different paths' information. Furthermore, we adopt Squeeze-and-Excitation (SE) with HP-GTC modules, which can alleviate the over-smoothing problem and enhance influential features. Finally, we apply a readout layer to generate the final graph embedding to estimate brain age and gender, and thoroughly evaluate the PH-BTN on the Baby Connectome Project (BCP) dataset. Experimental results demonstrate the superiority of PH-BTN over other state-of-the-art methods. The proposed PH-BTN offers a powerful tool to investigate and explore brain functional connectivity.

**Keywords:** Path-based heterogeneous network · Heterogeneous path graph transformer convolution · Brain functional connectivity analysis

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43993-3_32.


H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14227, pp. 328–337, 2023.
https://doi.org/10.1007/978-3-031-43993-3_32

## 1 Introduction

Brain functional network refers to the integrator of information exchange between different neurons, neuron clusters or brain regions. It can not only reveal the working mechanism and developmental changes of the brain [24], but also reflect anatomical connectivity of brain structure [12], making it a hot topic of neuroscience research in recent years. Current research [15,19,32] based on brain functional network mainly focus on two directions: brain physiological basis and brain diseases. Brain diseases are often associated with abnormal connections and have been shown to be related to physiological basis, especially age and sex [2,4,9]. In neuroscience, age and gender prediction based on brain functional network would be the basis for better studying brain diseases, and understanding and exploring the operating mechanisms of baby brains.

As a powerful neuroimaging tool, the resting-sate functional Magnetic Resonance Imaging (rs-fMRI) constructs brain functional networks by capturing the changes of blood oxygen level-dependent (BOLD) signals and computing their correlation between different regions of interest (ROIs) [5]. Owing to the benefits of non-invasive and high-resolution of rs-fMRI, resting-state functional connectivity (rs-FC) derived from BOLD signals is increasingly used to analyze brain age and gender [6]. Up to now, rs-FC analysis methods have primarily included correlation-based methods and graph-based approaches [26]. Compared with correlation-based methods, representing rs-FC data as a graph can preserve the natural topological properties of brain networks, where the nodes are defined as ROIs by an atlas, and the edges are calculated as pairwise correlations between ROIs. However, the assumptions of most existing graph-based methods are still far from the reality of the human brain with the following limitations:

***Ignoring Path Importance***. Some studies have indicated that connectivity is the core of the brain and no neuron is an island [1,21]. Similar as in the graph theory, these connections can be defined as paths in a brain graph. However, popular methods PR-GNN [18] and BrainGNN [17] mainly focused on node features, which ignored the significance of path features. Despite the fact that BrainNetCNN [14] built an edge-based brain network which can be viewed as a special path-based network, the edge-based description can only depict the direct connections, but cannot account for a variety of indirect connections between brain regions.

***Neglecting Heterogeneity***. Although BC-GCN [19] formulated brain network as a path-based graph, the graph was homogeneous with only one type of path. In fact, brain functional network is heterogeneous, as proved by abundant studies [7, 22,30].

***Overlooking Global Structures***. Attention mechanism can help the model focus on crucial connections. Inspired by this, GAT model [25] was employed to analyze brain networks in [31], but limited by graph convolution, it only considered the local structures of neighboring nodes. According to emerging research, global and local structures provide different views for brain network analysis [8,28], and global brain information concerned by Graph Transformer can further improve predictive performance [3,13].

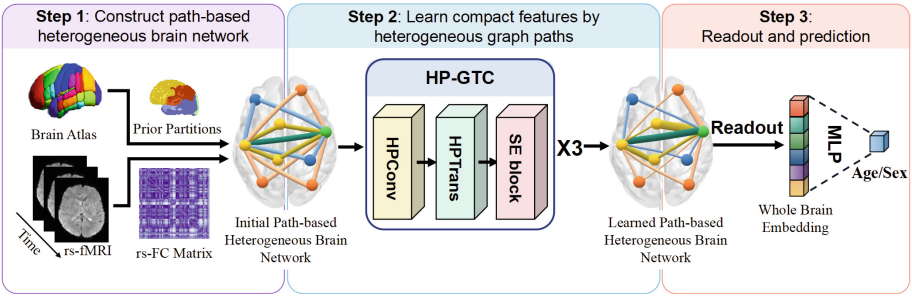


**Fig. 1.** An overview of PH-BTN for brain age or gender prediction. The color only represents type of paths and the thickness of edges/paths reflects the change of features.

To address these issues, we propose to construct a **P**ath-based **H**eterogeneous **B**rain **T**ransformer **N**etwork (PH-BTN) from rs-fMRI data, learn compact edge features by heterogeneous graph paths and finally readout for brain rs-FC analysis. The proposed PH-BTN compensates the graph neural network (GNN) with the capacity of modeling heterogeneous and global information of the brain graph by designing the **H**eterogeneous **P**ath **G**raph **T**ransformer **C**onvolution (HP-GTC) module with multi-path hypothesis. The major contributions of this work are highlighted below:

- We offer a new perspective on modeling the brain network as a heterogeneous graph with multiple types of path-based features under the prior knowledge of brain partitions, which takes into account the path significance and heterogeneity of the brain, and better simulates the brain network.
- We develop a novel Graph Transformer Network, namely PH-BTN, which is unprecedently able to encode path-based heterogeneous brain network with Transformer to enforce neighbors while incorporating global information. The core component in PH-BTN, namely HP-GTC module, is combined with heterogeneous graph convolution and attention, which can aggregate rich and crucial path information to generate compact brain representation. Furthermore, the Squeeze-and-Excitation (SE) block [11] adopted in HP-GTC module can alleviate the over-smoothing problem of GNN.
- We conduct extensive experiments on the Baby Connectome Project (BCP) dataset to verify the superiority of our proposed method compared with other state-of-the-art methods, and explore the age and gender relevance to the brain functional network.

## 2 Methodology

An overview of the proposed PH-BTN is illustrated in Fig. 1. Below, we first introduce the construction of the path-based heterogeneous brain network. Then, we focus on the HP-GTC module and elaborate on the novel design and its two components with SE block, i.e., heterogeneous graph path convolution and transformer layer. Finally, we briefly describe the readout and prediction stages.

### 2.1 Path-Based Heterogeneous Graph Generation

**Path-Based Heterogeneous Brain Network.** To retain the heterogeneity and path-based structure, we encode the brain functional network captured by rs-fMRI as a path-based heterogeneous graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{P})$, where $\mathcal{V} = \{v_i\}_{i=1}^{N} \in \mathbb{R}^N$ is the node set of size $N$ defined by ROIs on a specific brain atlas, $\mathcal{E} = [e_{ij}] \in \mathbb{R}^{N \times N \times D}$ is the edge set constructed by using the Pearson's correlation coefficient (PCC) between a sub-series of BOLD signals between nodes, with each edge $e_{ij}$ initialized with $D$-dimensional edge features $h_{ij}$, and $\mathcal{P}$ is the path set along with a path type mapping function $\boldsymbol{\psi} : \mathcal{P} \rightarrow \mathcal{R}$, where $\mathcal{R}$ denotes the graph path types, $|\mathcal{R}| \geq 2$.

**Graph Path.** In graph theory, a graph path $p$ is composed of a finite sequence of $n$ edges, where $n$ denotes that this path is a n-hop path $p^n$ . For example, a 2-hop path $p^2_{ikj}$ between node $v_i$ and $v_j$ can be represented as $p^2_{ikj} = \{e_{ik}, e_{kj}\}$, where $e_{ij}$ is the edge between $v_i$ and $v_j$ and $i \neq j \neq k$. In particular, the 0-hop path indicates the self-loop of the node. According to ablation results of multi-hops [20], we can see that graph paths with finite hops contain enough effective information. To simplify, we limit the highest hop to 2. Then, the multi-hop paths between node $v_i$ and $v_j$ can be defined as $P_{ij} := (P^0_{ij}, P^1_{ij}, P^2_{ij})$, where $P^n_{ij}$ means the set of all n-hop paths between node $v_i$ and $v_j$, $P^0_{ij} = \{e_{ij} | i = j\}$ and $P^1_{ij} = \{e_{ij} | i \neq j\}$. Particularly, when combining $P^0_{ij}$ and $P^1_{ij}$, we can gain the sequence $\{e_{ik}, e_{kj}\}$, where $i = k \neq j$ *or* $i \neq k = j$. This conjunctive sequence of $P^0_{ij}$ and $P^1_{ij}$ can be regarded as a special 2-hop paths $\{p_{ikj} | i = k \neq j \ or \ i \neq k = j\}$, where $k$ denotes the index of intermediate node. Thus, the multi-hop paths between node $v_i$ and $v_j$ can be recorded as $P_{ij} := P_{ikj} = \{p_{ikj} | \forall k \in N\}$.

**Heterogeneous Graph Path.** In order to obtain the type of paths, we introduce a brain partition (e.g., frontal, parietal, temporal, occipital and insular) as prior to define heterogeneous paths. It is intuitive to define the types of graph path by edges. However, with the increase of edges, the complexity of the path type definition will increase greatly. Thus, we directly regard the type of intermediate node $v_k$ as the type of graph path $p_{ikj}$. For instance, if $v_k$ in the frontal lobe, we have $\boldsymbol{\psi}(p_{ikj}) = frontal$. Considering the fact that the feature spaces of different types of paths are not completely irrelevant (as shown in Fig. 3-(a)), we have additionally defined all graph path as the base type, which can reflect the common information shared by different types of graph paths. For example, when using the above-mentioned partition as prior to define path type, we have $|\mathcal{R}| = |\{base, frontal, parietal, temporal, occipital, insular\}| = 6$.

### 2.2 HP-GTC: Heterogeneous Path Graph Transformer Convolution Module to Learn Compact Features

After brain network construction, we present the HP-GTC module to extract and aggregate specific and common features from different types of graph paths. As illustrated in Fig. 2, the HP-GTC module consists of two layers, which are **H**eterogeneous Graph **P**ath **Conv**olution (HPConv) layer and **H**eterogeneous Graph **P**ath **Trans**former (HPTrans) layer respectively.

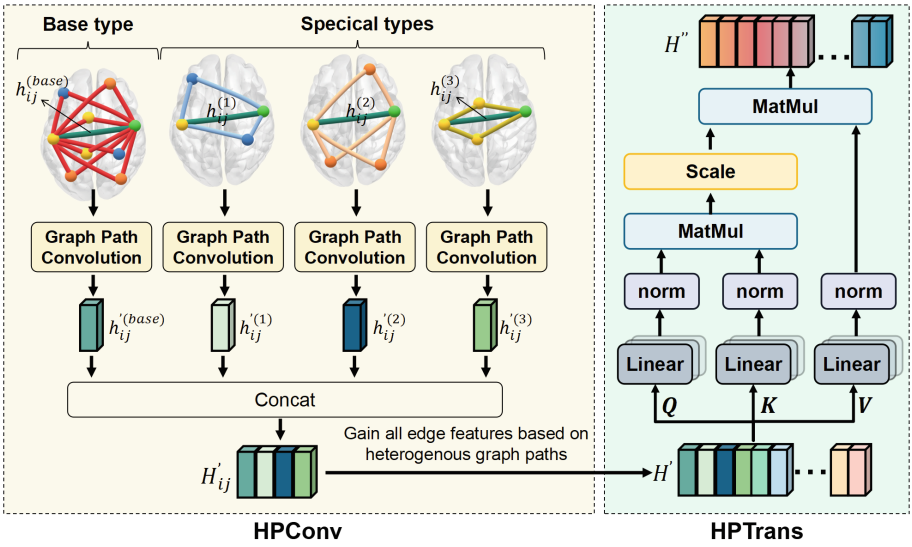


**Fig. 2.** Diagram for the details of HPConv and HPTrans layers for feature learning. For the sake of simplicity and clarity, we here only show the above four paths' types as an example, which does not mean that we only use four types.

**HPConv.** Considering that different heterogeneous paths have different distributions and contain different information, we design a novel graph convolution to learn the edge features of each path type independently. Under each path type, the brain network can be regarded as a homogeneous graph. Inspired by Li et al. [19], we utilize the same graph path convolution for each path type here. Within each path-based heterogeneous brain network $\mathcal{G}$, we define the following simple propagation model as HPConv layer for the forward-pass update of an edge denoted by $e_{ij}$ under the graph type $r \in \mathcal{R}$:

$$\boldsymbol{h}_{ij}^{'(r)} = \sum_{k:\ \psi(p_{ikj})=r} p_{ikj}^{(r)} \boldsymbol{w}_k^{(r)} = \sum_{k:\ \psi(p_{ikj})=r} \left( h_{ik}^{(r)} + h_{kj}^{(r)} \right) \boldsymbol{w}_k^{(r)} \tag{1}$$

where $\boldsymbol{h}_{ij}^{(r)} \in \mathbb{R}^d$ and $\boldsymbol{h}_{ij}^{'(r)} \in \mathbb{R}^{d'}$ are the input and output features of HPConv layer in the $l^{th}$ HP-GTC module under the graph type $r$, $d$ and $d'$ are features' dimension, $p_{ikj}^{(r)}$ is the path feature and $\boldsymbol{w}_k^{(r)}$ denotes the learnable parameters of transformation. Thus, we finally have the intermediate representation feature matrix $\boldsymbol{H}^{'(r)} = \left[ \boldsymbol{h}_{ij}^{'(r)} \right]$ under the graph path type $r$. Intuitively, the HPConv layer extracts edge features in specific and common path spaces, so as to ensure that the edge features learned in different specific feature spaces are independent and those learned in common spaces can retain global and shared information.

**HPTrans.** Note that different types of graph paths would have different impacts on a specific edge, and the same type of graph paths may similarly affect the specific edge. To model these characteristics, we propose an attention mechanism to capture these heterogeneous and common information to learn more effective edge features. Inspired by Kan et al. [13], to further incorporate global information and inconsistent contributions from different types of paths, we design a novel Graph Transformer layer to further aggregate output features $\boldsymbol{H}^{'} = \left[\boldsymbol{h}_{ij}^{'(r\in\mathcal{R})}\right]$ of HPConv layer for learning more compact edge features. The HPTrans layer is formulated as follows:

$$\boldsymbol{Q} = norm(\boldsymbol{H}^{'}\boldsymbol{W}_Q),\ \boldsymbol{K} = norm(\boldsymbol{H}^{'}\boldsymbol{W}_K),\ \boldsymbol{V} = norm(\boldsymbol{H}^{'}\boldsymbol{W}_V) \tag{2}$$

$$\boldsymbol{A} = (\boldsymbol{Q}\boldsymbol{K}^{\top})/\sqrt{|R|},\ \boldsymbol{H}^{''} = HPTrans(\boldsymbol{H}^{'}) = \boldsymbol{A}\boldsymbol{V} \tag{3}$$

where $\boldsymbol{H}^{''}$ denotes the output features, and the weight parameters $\boldsymbol{W_Q}$, $\boldsymbol{W_K}$, $\boldsymbol{W_V} \in \mathbb{R}^{d'\times\frac{d'}{2}}$ linearly map edge features to different feature spaces, and then adopt L2-normalized operator $norm(\cdot)$ to generate the corresponding representations $\boldsymbol{Q}$, $\boldsymbol{K}$, $\boldsymbol{V}$ for subsequent attention matrix $\boldsymbol{A}$ calculation and feature aggregation. For simplicity of illustration, in this paper, we only consider the single-head self-attention and assume $\boldsymbol{V} = \boldsymbol{H}^{'}$. The extension to the multi-head attention is straightforward, and we omit bias terms for simplicity.

To further enhance influential features and alleviate the over-smoothing problem of GNN, we adopt SE block in the HP-GTC module. Therefore, the final formulation of HP-GTC module can be represented as below:

$$\boldsymbol{H}^{l+1} = SE\left(HPTrans\left(HPConv\left(\boldsymbol{H}^{l}\right)\right)\right) \tag{4}$$

where $\boldsymbol{H}^{l}$ is the edge feature matrix of $l^{th}$ HP-GTC module.

### 2.3 Readout and Prediction

Lastly, inspired by Li et al. [19], a readout layer is adopted to transform the edge features learned by heterogeneous paths into the final graph embedding $\boldsymbol{H}_{G_m}$. Then $\boldsymbol{H}_{G_m}$ is sent to a multi-layer perceptron (MLP) to give the final prediction $\hat{y}_m$, where $G_m$ denotes the $m^{th}$ graph sampled in subject $s$. Specifically, the final prediction of subject $s$ is the average or weighted voting result of its' all samples during inferring.

## 3 Experimental Results

**Dataset and Implementation Details.** We validated our PH-BTN on the Baby Connectome Project (BCP) dataset [10] including 612 longitudinal rs-fMRI scans from 248 subjects (ages 6-811days, 106 boys vs. 142 girls). We here adopted Harvard-Oxford atlas (N = 112 ROIs) [23] and the mainstream brain functional partition as prior, where $|\mathcal{R}|$ =

**Table 1.** Brain age prediction and sex classification results (mean ± std) of all comparison methods on BCP dataset. (**bold**: best; underline: runner-up)

| Graph mode | | Method | Age Prediction | | Sex Classification | |
|---|---|---|---|---|---|---|
| | | | MAE(days) ↓ | PCC(%) ↑ | ACC(%) ↑ | F1(%) ↑ |
| Non-graph | | MLP | $82.16 \pm 0.46$ | $86.16 \pm 1.08$ | $74.57 \pm 5.51$ | $74.14 \pm 6.06$ |
| | | CNN | $89.90 \pm 1.72$ | $84.49 \pm 4.23$ | $76.62 \pm 7.28$ | $75.88 \pm 6.85$ |
| Homogeneous | Node-based | GCN [16] | $81.21 \pm 1.67$ | $86.27 \pm 4.12$ | $77.42 \pm 7.86$ | $75.62 \pm 9.29$ |
| | | PR-GNN [18] | $79.02 \pm 2.51$ | $87.91 \pm 5.65$ | $78.65 \pm 4.92$ | $78.29 \pm 4.87$ |
| | | BrainGNN [17] | $75.55 \pm 1.94$ | $88.69 \pm 4.54$ | $77.43 \pm 6.53$ | $76.65 \pm 7.03$ |
| | Edge-based | BrainNetCNN [14] | $70.57 \pm 1.25$ | $89.65 \pm 3.82$ | $79.03 \pm 6.21$ | $78.23 \pm 6.95$ |
| | Path-based | BC-GCN$_{SE}$ [19] | $\underline{61.40 \pm 1.02}$ | $\underline{90.32 \pm 3.64}$ | $\underline{79.43 \pm 5.82}$ | $\underline{78.63 \pm 5.69}$ |
| Heterogeneous | | PH-BTN | $\mathbf{58.91 \pm 1.64}$ | $\mathbf{92.56 \pm 2.16}$ | $\mathbf{81.87 \pm 5.06}$ | $\mathbf{81.50 \pm 5.18}$ |

$|\{base,\ frontal,\ parietal,\ temporal,\ occipital, insular\}| = 6$. For experiments, we evaluated PH-BTN and related competing models (i.e., four brain backbone models - BrainNetCNN, BrainGNN, PR-GNN and BC-GCN, and three classical deep learning models - MLP, CNN and GCN) through 10-fold cross-validation. We chose the classification accuracy (ACC) and weighted F1-score (F1) as evaluation metrics for sex classification, and evaluated the Mean Absolute Error (MAE) and Pearson's Correlation Coefficient (PCC) between predicted and ground truth ages. To guarantee training is fair, we used the identical parameters across all experimental baselines. See *Supplementary Materials* for more details of dataset and implementation.

**Performance Analysis.** As shown in Table 1, we first report the overall results of age prediction and sex classification achieved by the proposed PH-BTN and related alternative methods on the BCP dataset. From Table 1, we can see that the proposed PH-BTN outperformed alternative models both on age prediction with MAE of 58.91 days (7.32% of the whole age distribution) and 92.56% PCC, and sex classification with 81.87% ACC and 81.50% F1. Furthermore, we have following interesting observations when comparing these statistics. First, graph-based methods are superior to other non-graph methods when utilizing rs-FC features, which demonstrates the rationality and effectiveness of formulating the brain as a graph. Second, we found that the performance of node-based methods is worse than that of edge-based/path-based methods, especially at least 5.28 days in age prediction. This confirms the view that the connectivity is the core of the brain. Third, we also observed that brain networks with heterogeneous design have higher ACC and lower MAE than the homogeneous design, which implies that the heterogeneity of brain connections. All above results indicate that our formulation is more in line with brain functional network by considering brain graph topology, path information and connective heterogeneity.

**Ablation Studies.** As shown in *Supplementary Materials*, we have done several ablation experiments to further evaluate the effectiveness of the proposed method, e.g., different path types by prior or random division, whether to use

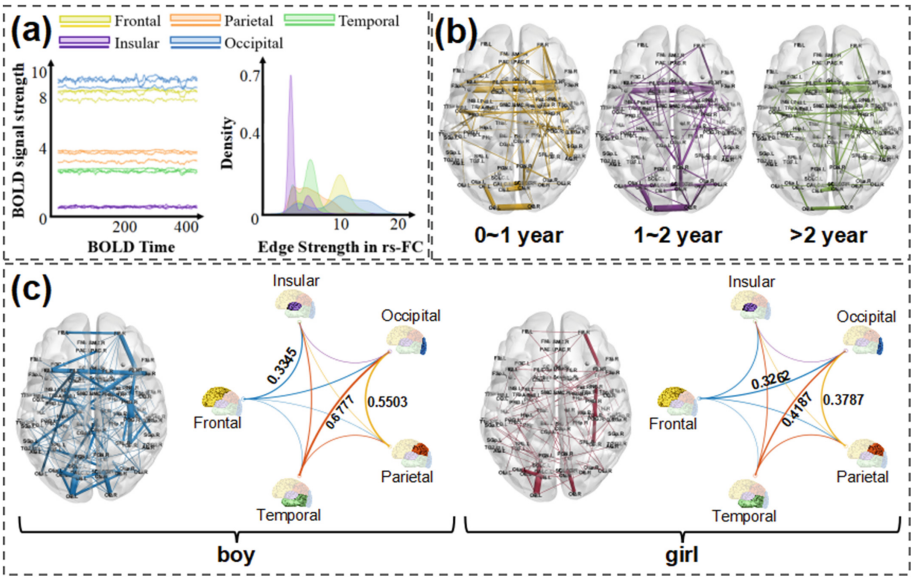


**Fig. 3.** (a) Heterogeneous analysis of five functional partitions before experiments. (b)–(c) Visualization of important connections of different age and sex learned by PH-BTN.

HPTrans or not, etc. After comparison, we can draw the following conclusions: (1) Reasonable path-type division is the key to learning brain heterogeneity. (2) Regardless of special types, the base type representing shared information is essential for brain functional connectivity analysis (w/o base loses at least 8.44 days). (3) The transformer-based mechanism can improve the performance (e.g., MAE 3.94 days) of PH-BTN in global and heterogeneous views.

**Visualization and Discussion.** In this section, we utilized the gradient backtracking method [19] to visualize brain functional networks learned by the proposed PH-BTN. For visualizing age-related connections, we divided all scans into three typical groups, i.e., 0–1 year, 1–2 year and >2 year. All of our visual findings illustrated in Fig. 3(b)–(c) are consistent with those of many research [7,27,29], which have showed that frontal, parietal and occipital lobes are deeply related to brain age and gender, and well correspond to our statistical analysis (Fig. 3(a)) of different brain regions in the original data. We refer the readers to the *Supplementary Materials* for additional results of visualization.

## 4 Conclusion

In this paper, we propose a novel network, namely PH-BTN, for encoding path-based heterogeneous brain networks for analyzing brain functional connectivity.

Different from most existing methods, our proposed model considers the path significance and heterogeneity by heterogeneous graph convolution, and incorporates global brain structure and key connections by Transformer mechanism. Experiments and visualization of age and gender prediction on the BCP dataset show the superiority and effectivity of our PH-BTN model. Moreover, the proposed PH-BTN offers a new way to understand neural development, explore sexual differences, and ultimately benefit neuroimaging research. In future work, we will extend and validate our methods on larger benchmark datasets.

**Acknowledgements.** This work was supported by Guangdong Key Laboratory of Human Digital Twin Technology (2022B1212010004) and Fundamental Research Funds for the Central Universities (2022ZYGXZR104). This work also utilized approaches developed by the efforts of the UNC/UMN Baby Connectome Project Consortium.

## References

1. Axer, M., Amunts, K.: Scale matters: the nested human connectome. Science **378**(6619), 500–504 (2022)
2. Bao, A.M., Swaab, D.F.: Sex differences in the brain, behavior, and neuropsychiatric disorders. Neuroscientist **16**(5), 550–565 (2010)
3. Cai, H., Gao, Y., Liu, M.: Graph transformer geometric learning of brain networks using multimodal MR images for brain age estimation. IEEE Trans. Med. Imaging **42**(2), 456–466 (2023)
4. Cole, J.H., Franke, K.: Predicting age using neuroimaging: innovative brain ageing biomarkers. Trends Neurosci. **40**(12), 681–690 (2017)
5. Friston, K.J.: Functional and effective connectivity in neuroimaging: a synthesis. Hum. Brain Mapp. **2**(1–2), 56–78 (1994)
6. Gadgil, S., Zhao, Q., Pfefferbaum, A., Sullivan, E.V., Adeli, E., Pohl, K.M.: Spatio-temporal graph convolution for resting-state fMRI analysis. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12267, pp. 528–538. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59728-3_52
7. Gao, W., Alcauter, S., Smith, J.K., Gilmore, J.H., Lin, W.: Development of human brain cortical network architecture during infancy. Brain Struct. Funct. **220**, 1173–1186 (2015)
8. He, S., Grant, P.E., Ou, Y.: Global-local transformer for brain age estimation. IEEE Trans. Med. Imaging **41**(1), 213–224 (2021)
9. Hou, Y., et al.: Ageing as a risk factor for neurodegenerative disease. Nat. Rev. Neurol. **15**(10), 565–581 (2019)
10. Howell, B.R., et al.: The UNC/UMN baby connectome project (BCP): an overview of the study design and protocol development. Neuroimage **185**, 891–905 (2019)
11. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7132–7141 (2018)
12. Jung, J., Cloutman, L.L., Binney, R.J., Ralph, M.A.L.: The structural connectivity of higher order association cortices reflects human functional brain networks. Cortex **97**, 221–239 (2017)
13. Kan, X., Dai, W., Cui, H., Zhang, Z., Guo, Y., Yang, C.: Brain network transformer. arXiv preprint arXiv:2210.06681 (2022)

14. Kawahara, J., et al.: Brainnetcnn: convolutional neural networks for brain networks; towards predicting neurodevelopment. Neuroimage **146**, 1038–1049 (2017)
15. Kim, B.H., Ye, J.C.: Understanding graph isomorphism network for rs-fMRI functional connectivity analysis. Front. Neurosci. **14**, 630 (2020)
16. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
17. Li, X., et al.: BrainGNN: interpretable brain graph neural network for fMRI analysis. Med. Image Anal. **74**, 102233 (2021)
18. Li, X., et al.: Pooling regularized graph neural network for fMRI biomarker analysis. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12267, pp. 625–635. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59728-3_61
19. Li, Y., et al.: Brain connectivity based graph convolutional networks and its application to infant age prediction. IEEE Trans. Med. Imaging **41**(10), 2764–2776 (2022)
20. Nikolentzos, G., Dasoulas, G., Vazirgiannis, M.: K-hop graph neural networks. Neural Netw. **130**, 195–205 (2020)
21. Thiebaut de Schotten, M., Forkel, S.J.: The emergent properties of the connected brain. Science **378**(6619), 505–510 (2022)
22. Shi, G., Zhu, Y., Liu, W., Yao, Q., Li, X.: Heterogeneous graph-based multimodal brain network learning. arXiv e-prints pp. arXiv-2110 (2021)
23. Smith, S.M., et al.: Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage **23**, S208–S219 (2004)
24. Van Den Heuvel, M.P., Pol, H.E.H.: Exploring the brain network: a review on resting-state fMRI functional connectivity. Eur. Neuropsychopharmacol. **20**(8), 519–534 (2010)
25. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y., et al.: Graph attention networks. Stat **1050**(20), 10–48550 (2017)
26. Wang, L., Li, K., Hu, X.P.: Graph convolutional network for fMRI analysis based on connectivity neighborhood. Netw. Neurosci. **5**(1), 83–95 (2021)
27. Weis, S., Patil, K.R., Hoffstaedter, F., Nostro, A., Yeo, B.T., Eickhoff, S.B.: Sex classification by resting state brain connectivity. Cereb. Cortex **30**(2), 824–835 (2020)
28. Wen, X., Wang, R., Yin, W., Lin, W., Zhang, H., Shen, D.: Development of dynamic functional architecture during early infancy. Cereb. Cortex **30**(11), 5626–5638 (2020)
29. Wu, K., Taki, Y., Sato, K., Hashizume, H., Sassa, Y., et al.: Topological organization of functional brain networks in healthy children: differences in relation to age, sex, and intelligence. PLoS ONE **8**(2), e55347 (2013)
30. Yao, D., Yang, E., Sun, L., Sui, J., Liu, M.: Integrating multimodal MRIs for adult ADHD identification with heterogeneous graph attention convolutional network. In: Rekik, I., Adeli, E., Park, S.H., Schnabel, J. (eds.) PRIME 2021. LNCS, vol. 12928, pp. 157–167. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87602-9_15
31. Yin, W., Li, L., Wu, F.X.: A graph attention neural network for diagnosing ASD with fMRI data. In: 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1131–1136. IEEE (2021)
32. Zhang, H., et al.: Classification of brain disorders in rs-fMRI via local-to-global graph neural networks. IEEE Trans. Med. Imaging **42**(2), 444–455 (2023)