Prediction of Infant Cognitive Development with Cortical Surface-Based Multimodal Learning

Jiale Cheng[1,2], Xin Zhang[1,3](✉), Fenqiang Zhao[2], Zhengwang Wu[2], Xinrui Yuan[2], Li Wang[2], Weili Lin[2], and Gang Li[2](✉)

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

Abstract. Exploring the relationship between the cognitive ability and infant cortical structural and functional development is critically important to advance our understanding of early brain development, which, however, is very challenging due to the complex and dynamic brain development in early postnatal stages. Conventional approaches typically use either the structural MRI or resting-state functional MRI and rely on the region-level features or inter-region connectivity features after cortical parcellation for predicting cognitive scores. However, these methods have two major issues: 1) *spatial information loss*, which discards the critical fine-grained spatial patterns containing rich information related to cognitive development; 2) *modality information loss*, which ignores the complementary information and the interaction between the structural and functional images. To address these issues, we unprecedentedly invent a novel framework, namely cortical surface-based multimodal learning framework (CSML), to leverage fine-grained multimodal features for cognition development prediction. First, we introduce the fine-grained surface-based data representation to capture spatially detailed structural and functional information. Then, a dual-branch network is proposed to extract the discriminative features for each modality respectively and further captures the modality-shared and complementary information with a disentanglement strategy. Finally, an age-guided cognition prediction module is developed based on the prior that the cognition develops along with age. We validate our method on an infant multimodal MRI dataset with 318 scans. Compared to state-of-the-art methods, our method consistently achieves superior performances, and for the first time suggests crucial regions and features for cognition development hidden in the fine-grained spatial details of cortical structure and function.

Keywords: Cognition Prediction · Multimodality · rs-fMRI · sMRI

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14221, pp. 618–627, 2023.
https://doi.org/10.1007/978-3-031-43895-0_58

1 Introduction

Predictive modeling of the individual-level cognitive development during infancy is of great importance in advancing our understanding of the subject-specific relationship between the cognitive ability and early brain structural and functional development and their underlying neural mechanisms. It is also critical for early identifying cognitive delays and developing more effectively and timely personalized therapeutic interventions for at-risk infants. However, this is a very challenging task due to the complex and rapid development of brain structure, function and cognition during the first years of life [1, 2].

Recently, a few methods have been explored for predicting infant cognition using either resting-state functional MRI (rs-fMRI) [2–4] or structural MRI (sMRI) [5–7]. Although encouraging preliminary results have been achieved, two unaddressed major issues hinder the precise prediction of the individual-level cognitive development during infancy. 1) *Spatial information loss*: Previous works [3–8] typically rely on region-level features or inter-region connectivity features after parcellation of the brain cortex into a set of regions. Consequently, these features largely ignore fine-grained spatial patterns on cortical surfaces, which encode subject-specific rich information critical for cognitive prediction. 2) *Modality-information loss*: Previous methods use either functional features or structural features, and thus the complementary information between them and their underlying relationship are not leveraged for cognition development. Indeed, it is believed that the spontaneous neuronal activity is related to the intrinsic human brain functional organizations supported by the underlying structural substrates [2], which gives emphasis to understanding the underlying individual structure-functional profile during infancy. Therefore, an effective unified framework that can automatically learn the complementary and spatially fine-grained information from structural and functional data for cognition development prediction is critically desired.

To address the above issues, we propose a novel cortical surface-based multimodal learning framework (**CSML**), to enable learning of the fine-grained spatial patterns and complementary information from structural and functional MRI data for precise prediction of the individual-level cognitive development. Specifically, 1) to learn detailed spatial patterns of both functional connectivity and structural information, we propose to leverage the strong feature learning and representation ability of spherical surface networks [9] to automatically extract task-related features on cortical surfaces. 2) To effectively fuse structural and functional information, we propose a dual-branch surface network to simultaneously extract structural morphologic features and functional connectivity features on cortical surfaces, and further fuse their complementary information in a feature disentanglement module. 3) To enable precise prediction of cognitive outcome, we leverage the prior knowledge that the cognition function develops with age growing by jointly predicting age and cognition scales. To our best knowledge, this is the first work to leverage the multimodal, fine-grained spatial information on cortical surface explicitly for cognition development prediction. The experimental results based on a longitudinal infant dataset not only validate the superiority of our proposed model but also imply the tight association between the individual cognition development and the fine-grained cortical information.

2 Method

In this section, we present the details of CSML (Fig. 1), including three steps: 1) surface-based fine-grained information representation (Fig. 1(a)); 2) modality-specific information learning (Fig. 1(b)); and 3) multi-modality information fusion (Fig. 1(c)).

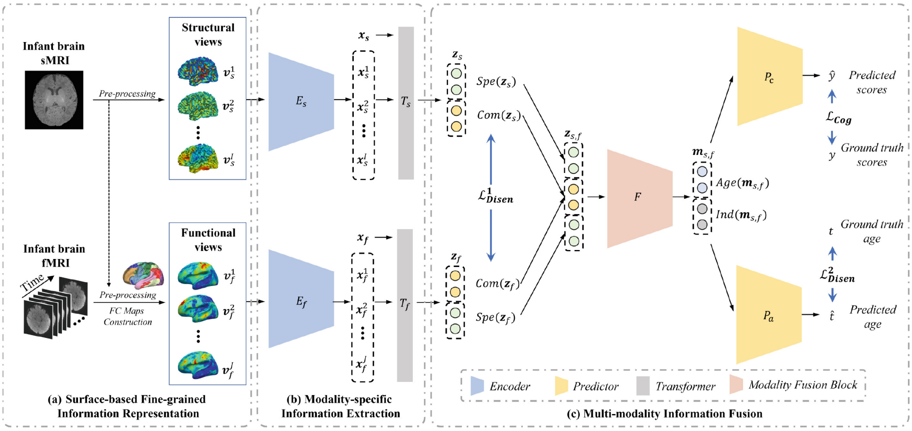

Fig. 1. Overview of our framework for cortical surface-based multimodal fine-grained information learning. Given the sMRI and fMRI of an infant, its structural and functional feature representations z_s and z_f are first extracted. Then, the modality shared ($Com(z)$) and specific ($Spe(z)$) information are disentangled and further fused by a modality fusion block F. After that, we constrain the fused latent variable $\boldsymbol{m}_{s,f}$ to be age-irrelevant by the age predictor P_a, and finally obtain the predicted cognitive scores from the predictor P_c.

2.1 Surface-Based Fine-Grained Information Representation

The input of the network framework consists of two branches for encoding cortical structural information and functional connectivity information, respectively. To integrate multi-modal MRI data together for cognition development prediction, we map all modality data to a common space, i.e., the cortical surface registered to UNC 4D infant surface atlas [10] and further resampled with 40,962 vertices, following the well-established pipelines [11–15]. To capture the spatially fine-grained information in structural MRI, the *structural branch* contains a set of surface maps of biologically meaningful cortical properties, including cortical thickness, surface area, cortical volume, sulcal depth, mean curvature, and average convexity. To preserve the fine-grained spatial patterns of functional connectivity, we leverage an infant-dedicated cortical functional parcellation map [15]. Specifically, for each parcel, we first calculate the Pearson's correlation coefficient between the averaged functional time series of all vertices within this parcel and the functional time series of each cortical vertex to build the parcel-specific cortical functional connectivity (FC) map and then perform Fisher's r-to-z transformation. Finally, we use these cortical FC maps from all parcels, which characterize rich spatially detailed FC information, as the input of the *functional branch*.

2.2 Modality-Specific Encoder

For the multi-modality input, we employ two modality-specific encoders E_s and E_f to describe its feature representation, respectively. To be specific, we regard each modality comprised of multiple feature channels, while each channel could be interpreted as an observation of the data from a certain view. Therefore, we process each view separately as $\boldsymbol{x}_s^i = E_s\left(\boldsymbol{v}_s^i\right), \boldsymbol{x}_f^j = E_f\left(\boldsymbol{v}_f^j\right)$, where $i \epsilon [1, I]$, I is the number of morphological features we used; $j \epsilon [1, J]$, J is the number of parcels we used in building FC maps. We implement E_s and E_f as the Spherical Res-Net [9, 16, 17], which is composed of stacks of spherical convolutional layers and spherical pooling layers to extract the fine-grained spatial patterns and generates the view-related feature representations of $\boldsymbol{v}_s^i$ and $\boldsymbol{v}_f^j$. Considering the different number of views in each modality, two Transformer layers [18] T_s and T_f are then adopted to fuse the multi-view feature representations for each modality separately. Herein, following the previous work [19], we prepend two learnable embeddings $\boldsymbol{x}_s$ and $\boldsymbol{x}_f$ as the first token for the sequences of view-related feature representations $\{\boldsymbol{v}_s^i | i \epsilon [1, I]\}$ and $\{\boldsymbol{v}_f^j | j \epsilon [1, J]\}$, respectively. Within the Transformer layers, for the structural-related features, $\boldsymbol{x}_s$ interact with and fuse the view-related feature representation $\{\boldsymbol{v}_s^i | i \epsilon [1, I]\}$ through the self-attention mechanism as follows,

$$A_s^i = Q_s(\boldsymbol{x}_s) K_s\left(\boldsymbol{v}_s^i\right)^T / const, \tag{1}$$

$$\widetilde{\boldsymbol{x}}_s = \boldsymbol{x}_s + \sum_{i=1}^{I} softmax\left(A_s^i\right) U_s(\boldsymbol{v}_s^i), \tag{2}$$

$$z_s = \widetilde{\boldsymbol{x}}_s + W_s(\widetilde{\boldsymbol{x}}_s), \tag{3}$$

where z_s is the aggregated representation for the structural data, $Q_s(\cdot)$, $K_s(\cdot)$, $U_s(\cdot)$, and $W_s(\cdot)$ are four multi-layer perceptrons (*MLP*), *const* is a constant for normalization. Similarly, we can obtain the functional-related variable z_f by feeding $\boldsymbol{x}_f$ and the functional view-related representations $\{\boldsymbol{v}_f^j | j \epsilon [1, J]\}$ into T_f.

2.3 Modality-Fusion Block

To better learn the complementary information between the two modalities, we further decompose the modality-specific latent variables z_s and z_f into two parts: $Com(z_n)$ and $Spe(z_n)$, where $n \epsilon \{s, f\}$, standing for the structure (s) and function (f) related variables, respectively. $Com(z_n)$ is the common code representing the shared information among modalities, while $Spe(z_n)$ is the specific code representing the complementary information that differentiates one modality from the other. The basic requirements of this disentanglement are: (1) The concatenation of $Com(z_n)$ and $Spe(z_n)$ equals z_n; (2) $Com(z_s)$ and $Com(z_f)$ should be as similar as possible; (3) $Spe(z_s)$ differs from $Spe(z_f)$ as much as possible. Accordingly, $\mathcal{L}_{Disen}^1$ is defined as:

$$\mathcal{L}_{Disen}^1 = \mathcal{L}_{Disen}^{Com} / \mathcal{L}_{Disen}^{Spe}, \tag{4}$$

$$\mathcal{L}_{Disen}^{Com} = ||Com(z_s) - Com(z_f)||_2, \tag{5}$$

$$\mathcal{L}_{Disen}^{Spe} = ||Spe(z_s) - Spe(z_f)||_2. \tag{6}$$

Since the latent variable of each modality has been disentangled into $Com(z_n)$ and $Spe(z_n)$, the combined information is formed as the concatenation of the common code and specific codes as follows: $z_{s,f} = \big(Spe(z_s), Common, Spe(z_f)\big)$, where $Common = 0.5(Com(z_s) + Com(z_f))$.

2.4 Cognitive Scores Prediction

Given the combined multimodal information $z_{s,f}$, it is intuitive to regress the cognitive scores directly. However, considering that cognitive functions develop rapidly during the first years of life [1], the regressor would be prone to learn the age-related information instead and thus cannot differentiate the individualized development discrepancy between subjects within the same age group. Therefore, we fuse the combined multimodal information through a MLP F as follows, $\boldsymbol{m}_{s,f} = F(z_{s,f})$, and further disentangle the age-related variance $Age(\boldsymbol{m}_{s,f})$ and the individual-related invariance $Ind(\boldsymbol{m}_{s,f})$ from $\boldsymbol{m}_{s,f}$ to precisely evaluate the cognition development level. The basic requirements of this disentanglement are: (1) The concatenation of $Age(\boldsymbol{m}_{s,f})$ and $Ind(\boldsymbol{m}_{s,f})$ equals $\boldsymbol{m}_{s,f}$; (2) $Age(\boldsymbol{m}_{s,f})$ is capable of age estimation through an age predictor P_a; (3) $Ind(\boldsymbol{m}_{s,f})$ is incapable of age estimation through P_a. Accordingly, $\mathcal{L}_{Disen}^2$ is defined as:

$$\mathcal{L}_{Disen}^2 = \mathcal{L}_{Disen}^{Age} - \mathcal{L}_{Disen}^{Ind}, \tag{7}$$

$$\mathcal{L}_{Disen}^{Age} = |t - P_a(Age(\boldsymbol{m}_{s,f}))|, \tag{8}$$

$$\mathcal{L}_{Disen}^{Ind} = |t - P_a(Ind(\boldsymbol{m}_{s,f}))|, \tag{9}$$

where t is the ground truth of age. Then, we can use the identity-related features $Ind(\boldsymbol{m}_{s,f})$ containing subject-specific structure-function profile to predict the cognitive scores through a cognitive score predictor P_c under the guidance of the corresponding age feature $Age(\boldsymbol{m}_{s,f})$. The loss function to train P_c is defined as:

$$\mathcal{L}_{Cog} = |y - P_c(Ind\big(\boldsymbol{m}_{s,f}\big), Age(\boldsymbol{m}_{s,f}))|, \tag{10}$$

where y is the ground truth of cognitive scores. Specifically, we implement P_a and P_c as two sets of *MLP*. Finally, the overall objective function to optimize the neural network is written as:

$$L = \lambda_1 L_{Disen}^1 + \lambda_2 L_{Disen}^2 + L_{Cog}, \tag{11}$$

where λ_1 and λ_2 are trade-off parameters to balance the multiple losses.

3 Experiments

3.1 Dataset

We verified the effectiveness of the proposed CSML model for infant cognition development prediction on a public high-resolution dataset including 318 pairs of sMRI and rs-fMRI scans acquired at different ages ranging from 88 to 1040 days in the UNC/UMN Baby Connectome Project [20]. All structural and functional MR images were preprocessed following state-of-the-art infant-tailored pipelines [11–15]. Cortical surfaces were reconstructed and aligned onto the public UNC 4D infant surface atlas [10, 11]. For each cortical vertex on the middle cortical surface, its representative fMRI time-series was extracted [13–15]. An infant-dedicated fine-grained functional parcellation map [15] with 432 cortical ROIs per hemisphere in UNC 4D infant surface atlas was warped onto each individual cortical surface.

To quantify the cognition development level of each participant, four Mullen cognitive scores [21] were collected at their corresponding scan ages, i.e., Visual Receptive Scale (VRS), Fine Motor Scale (FMS), Receptive Language Scale (RLS), and Expressive Language Scale (ELS). These four cognitive scales were respectively normalized into the [0, 1] range using the minimum and maximum values for the training efficiency.

3.2 Experimental Settings

In order to validate our methods, a 5-fold cross-validation strategy is employed, and each fold consists of 190 training samples, 64 validating samples, and 64 testing samples. To quantitatively evaluate the performance, the Pearson's correlation coefficient (PCC) and root mean square error (RMSE) between the ground truth and predicted values were calculated. In the testing phase, the mean and standard deviation of the 5-fold results were reported.

The encoders E_s and E_f in CSML constitutes 5 Res-blocks with the dimensions of $\{32, 32, 64, 64, 128\}$, respectively. The modality fusion block F, age predictor P_a, and cognitive score predictor P_c were designed as two-layer MLP with the ReLU activation function and the dimension of $\{192, 128\}$, $\{64, 1\}$, and $\{128, 1\}$, respectively. We implemented the model with PyTorch and used Adam as optimizer with the weight decay of 10^{-4} and the learning rate cyclically tuned within $[10^{-6}, 10^{-3}]$. The batch size was set to 1. The maximum training epoch is 500. After comparison, we empirically set the hyperparameters as λ_1=0.05 and λ_2=0.01.

3.3 Results

We first show the results of some ablated models of our method in Table 1, where *w/o Structure* and *w/o Function* denote for the variants using functional and structural features only. *w/o Age* denotes the variant with single task of cognition prediction. It can be observed that, the overall performance on four cognitive tasks has been extensively improved by jointly leveraging the structural and functional information. The disentanglement mechanism successfully separates the shared and complementary information amongst modalities and further removes the redundancy with the loss $\mathcal{L}^1_{Disen}$. Moreover, the joint age prediction and cognitive estimation also brings further improvement by differentiating the age-related and identity-related variables with $\mathcal{L}^2_{Disen}$. The scatter plots of predicted cognitive scores in five testing folds are depicted in Fig. 2(a), demonstrating that the scores are well predicted.

We also comprehensively compared with various traditional and state-of-the-art functional connectivity-based methods, including KNN, random forest (RF), SVR, gaussian process regression (GPR), GCN [22], GAT [23], and UniMP [24]. As shown in Table 2, our algorithm outperforms the previous methods by a large margin. Of note, the proposed method demonstrates better performance even with the functional information only, which highlights the importance to preserve the fined-grained FC information.

Table 1. The impact of each component of CSML on the prediction performance (in terms of PCC). * indicates statistically significantly better results than other methods with p-value < 0.05.

Components	VRS	FMS	RLS	ELS	Average
w/o Age	0.768 ± 0.034	0.819 ± 0.023	0.789 ± 0.032	0.745 ± 0.034	0.780
w/o Structure	0.748 ± 0.050	0.814 ± 0.037	0.790 ± 0.042	0.711 ± 0.039	0.757
w/o Function	0.771 ± 0.038	0.831 ± 0.013	0.793 ± 0.028	0.737 ± 0.052	0.782
w/o $\mathcal{L}^1_{Disen}$	0.791 ± 0.015	0.837 ± 0.018	0.814 ± 0.018	0.789 ± 0.021	0.808
w/o $\mathcal{L}^2_{Disen}$	0.820 ± 0.045	0.858 ± 0.016	0.848 ± 0.016	0.824 ± 0.019	0.838
Proposed	**0.855 ± 0.026***	**0.874 ± 0.030***	**0.873 ± 0.008***	**0.852 ± 0.009***	**0.864***

Table 2. Performance comparison of different methods (in terms of RMSE and PCC). The averaged values among four tasks were provided. * indicates statistically significantly better results than other methods with p-value < 0.05.

Methods		RMSE	PCC
Machine Learning-based Methods	*KNN*	0.1421 ± 0.0144	0.5698 ± 0.0751
	RF	0.1353 ± 0.0060	0.6705 ± 0.1354
	GPR	0.1175 ± 0.0127	0.7320 ± 0.0596
	SVR	0.1284 ± 0.0069	0.7355 ± 0.0192
Graph Convolution-based Methods	*GCN*	0.1382 ± 0.0175	0.6234 ± 0.0581
	GAT	0.1208 ± 0.0228	0.7001 ± 0.1327
	UniMP	0.1246 ± 0.0149	0.7073 ± 0.0249
Proposed		**0.0915 ± 0.0091***	**0.8635 ± 0.0066***

Additionally, based on our proposed model CSML, the prediction accuracy of infant cognition development is over 0.85 on average, suggesting that the model may observe plausible biomarkers for cognition development during infancy. Based on the well-trained models, we explored the explainability and interpretability of the proposed method by investigating the weights of the Transformers. Since the Transformer layers T_s and T_f fuse the multi-view representations $\boldsymbol{v}_s^i$ and $\boldsymbol{v}_f^j$ into z_s and z_f for further cognitive prediction, by analyzing the attention value A_n^i of each view $\boldsymbol{v}_n^i$ in the Transformer, we can infer which regions for functional data and which morphological features for structural data are more important for cognition prediction. The results are shown in Fig. 2 (b) and Fig. 2 (c), in line with the reports in related studies to some extent [25–28], demonstrating the scientific value of our method. For example, the left lateral prefrontal cortex involved in higher executive functions [25, 26] demonstrates high importance in functional data. Moreover, previous researchers [27, 28] have observed the close relevance between the visual cortex and the early cognitive process, which also confirms the result of our method.

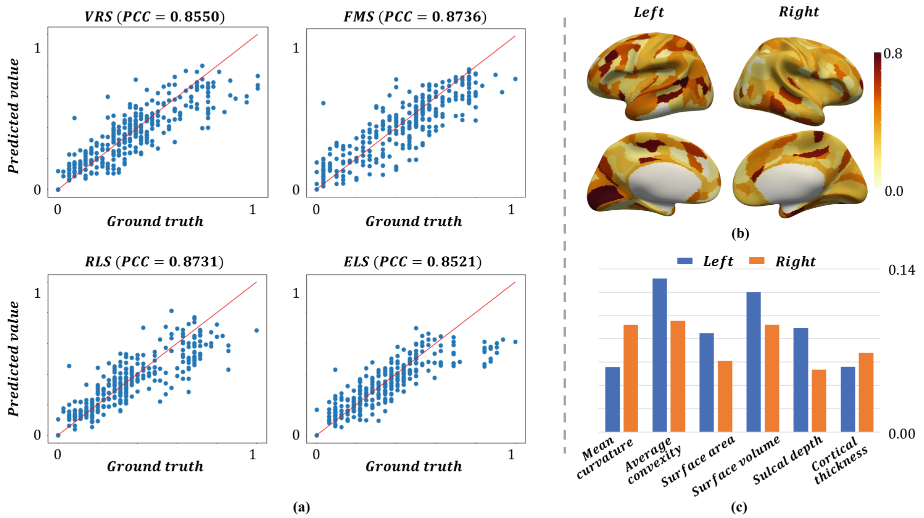

Fig. 2. The illustrations of (a) the predicted values distribution for four cognitive tasks, and the importance distribution of (b) each cortical regions and (c) each morphological feature on both hemispheres.

4 Conclusion

In this work, we develop an innovative cortical surface-based multimodal learning framework (CSML) to address the infant cognition prediction problem. Specifically, we unprecedentedly propose to explicitly leverage the surface-based feature representations to preserve the fine-grained, spatially detailed multimodal information for cognition prediction. In addition, by disentangling the modality-shared and complementary information, our model successfully captures the individualized cognition development patterns underlying the dramatic brain development. With its superior performance compared to state-of-the-art methods, our proposed CSML suggests that the informative clues for brain-cognitive relationship are hidden in the multimodal fine-grained details and validates itself as a potentially powerful framework for simultaneously learning effective representations from sMRI and rs-fMRI data.

Acknowledgements. The work of Gang Li was supported in part by NIH grants (MH116225, MH117943, MH127544, and MH123202). The work of Li Wang was supported by NIH grant (MH117943). This work also utilizes approaches developed by an NIH grant (1U01MH110274) and the efforts of the UNC/UMN Baby Connectome Project Consortium.

References

1. Gao, W., et al.: Functional network development during the first year: relative sequence and socioeconomic correlations. Cereb. Cortex **25**(9), 2919–2928 (2015)
2. Zhang, H., Shen, D., Lin, W.: Resting-state functional MRI studies on infant brains: a decade of gap-filling efforts. Neuroimage **185**, 664–684 (2019)
3. Keunen, K., Counsell, S.J., Benders, M.J.: The emergence of functional architecture during early brain development. Neuroimage **160**, 2–14 (2017)
4. Smyser, C.D., Snyder, A.Z., Neil, J.J.: Functional connectivity MRI in infants: exploration of the functional organization of the developing brain. Neuroimage **56**(3), 1437–1452 (2011)
5. Cheng, J., et al.: Path signature neural network of cortical features for prediction of infant cognitive scores. IEEE Trans. Med. Imaging **41**(7), 1665–1676 (2021)
6. Adeli, E., et al.: Multi-task prediction of infant cognitive scores from longitudinal incomplete neuroimaging data. Neuroimage **185**, 783–792 (2019)
7. Zhang, C., et al.: Infant brain development prediction with latent partial multi-view representation learning. IEEE Trans. Med. Imaging **38**(4), 909–918 (2018)
8. Hu, D., et al.: Existence of functional connectome fingerprint during infancy and its stability over months. J. Neurosci. **42**(3), 377–389 (2022)
9. Zhao, F., et al.: Spherical deformable u-net: application to cortical surface parcellation and development prediction. IEEE Trans. Med. Imaging **40**(4), 1217–1228 (2021)
10. Wu, Z., et al.: Construction of 4D infant cortical surface atlases with sharp folding patterns via spherical patch-based group-wise sparse representation. Hum. Brain Mapp. **40**(13), 3860–3880 (2019)
11. Li, G., et al.: Construction of 4D high-definition cortical surface atlases of infants: Methods and applications. Med. Image Anal. **25**(1), 22–36 (2015)
12. Li, G., et al.: Computational neuroanatomy of baby brains: a review. Neuroimage **185**, 906–925 (2019)
13. Li, G., et al.: Measuring the dynamic longitudinal cortex development in infants by reconstruction of temporally consistent cortical surfaces. Neuroimage **90**, 266–279 (2014)
14. Wang, L., Wu, Z., Chen, L., Sun, Y., Lin, W., Li, G.: iBEAT V2. 0: a multisite-applicable, deep learning-based pipeline for infant cerebral cortical surface reconstruction. Nat. Protoc. **18**(5), 1488–1509 (2023)
15. Wang, F., et al.: Fine-grained functional parcellation maps of the infant cerebral cortex. eLife (2023)
16. He, K., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
17. Zhao, F., et al.: Harmonization of infant cortical thickness using surface-to-surface cycle-consistent adversarial networks. In: Shen, Dinggang, et al. (eds.) MICCAI 2019. LNCS, vol. 11767, pp. 475–483. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32251-9_52

18. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, vol. 30, pp. 6000-6010 (2017)
19. Dosovitskiy, A., et al.: An image is worth 16x16 words: transformers for image recognition at scale. In International Conference on Learning Representation (2021)
20. 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)
21. Mullen, E.M.: Mullen scales of early learning. AGS Circle Pines, MN (1995)
22. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
23. Veličković, P., et al.: Graph attention networks. arXiv preprint arXiv:1710.10903 (2017)
24. Shi, Y., et al.: Masked label prediction: Unified message passing model for semi-supervised classification. arXiv preprint arXiv:2009.03509 (2020)
25. Fuster, J.M.: Frontal lobe and cognitive development. J. Neurocytol. **31**(3), 373–385 (2002)
26. Kolk, S.M., Rakic, P.: Development of prefrontal cortex. Neuropsychopharmacology **47**(1), 41–57 (2022)
27. Roelfsema, P.R., de Lange, F.P.: Early visual cortex as a multiscale cognitive blackboard. Ann. Rev. Vis. Sci. **2**, 131–151 (2016)
28. Albers, A.M., et al.: Shared representations for working memory and mental imagery in early visual cortex. Curr. Biol. **23**(15), 1427–1431 (2013)