Weakly Supervised Cerebellar Cortical Surface Parcellation with Self-Visual Representation Learning

Zhengwang Wu[1], Jiale Cheng[1], Fenqiang Zhao[1], Ya Wang[1], Yue Sun[1], Dajiang Zhu[2], Tianming Liu[3], Valerie Jewells[1], Weili Lin[1], Li Wang[1], and Gang Li[1](✉)

[1] Department of Radiology and Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
gang_li@med.unc.edu

[2] Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, TX, USA

[3] Department of Computer Science, University of Georgia, Athens, GA, USA

Abstract. The cerebellum (i.e., little brain) plays an important role in motion and balances control abilities, despite its much smaller size and deeper sulci compared to the cerebrum. Previous cerebellum studies mainly relied on and focused on conventional volumetric analysis, which ignores the extremely deep and highly convoluted nature of the cerebellar cortex. To better reveal localized functional and structural changes, we propose cortical surface-based analysis of the cerebellar cortex. Specifically, we first reconstruct the cerebellar cortical surfaces to represent and characterize the highly folded cerebellar cortex in a geometrically accurate and topologically correct manner. Then, we propose a novel method to automatically parcellate the cerebellar cortical surface into anatomically meaningful regions by a weakly supervised graph convolutional neural network. Instead of relying on registration or requiring mapping the cerebellar surface to a sphere, which are either inaccurate or have large geometric distortions due to the deep cerebellar sulci, our learning-based model directly deals with the original cerebellar cortical surface by decomposing this challenging task into two steps. First, we learn the effective representation of the cerebellar cortical surface patches with a contrastive self-learning framework. Then, we map the learned representations to parcellation labels. We have validated our method using data from the Baby Connectome Project and the experimental results demonstrate its superior effectiveness and accuracy, compared to existing methods.

Keywords: Cerebellar Cortex Parcellation · Representation Learning

1 Introduction

Parcellating the cerebellum (i.e., little brain) into neurobiologically meaningful regions of interest (ROIs) plays an important role in both structural and functional analysis [1, 2]. Therefore, automatic cerebellar parcellation methods are highly demanded to facilitate
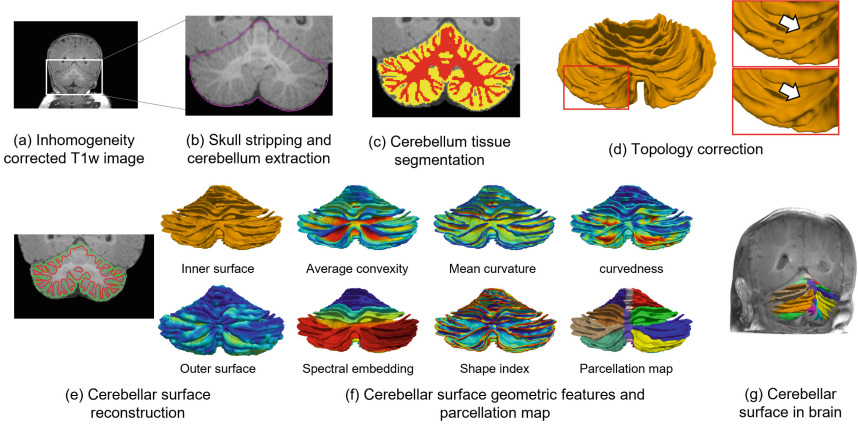
the analysis of increasingly larger cerebellum imaging datasets. Compared with the highly folded *cerebral* cortex, the *cerebellar* cortex has a more complex shape with even more tightly convoluted and much thinner cortical folds. Hence, few algorithms have been proposed for automatic cerebellar cortex parcellation.

Most existing methods mainly conduct the cerebellar cortex parcellation in the Euclidean space [3]. Generally, they treat the cerebellar cortex as a 3D volume, and then either use the traditional multi-atlas registration and label-fusion based methods or directly learn the cerebellar cortex parcellation model in the volumetric space, without reconstructing the cerebellar cortical surfaces and thus ignoring the topological and geometric properties of the cerebellar cortex. These methods are thus inappropriate for accurate and anatomically meaningful cerebellar cortex analysis, as the neighboring relationship in 3D space used in these methods cannot correctly reflect the actual cytoarchitectural neighborship. Specifically, two spatially neighboring points in the cerebellum volume do not necessarily mean they are neighboring in cytoarchitecture and geodesic. They could sit in completely different gyri or sulci, due to the extremely convoluted structure of the cerebellar cortex. This can cause severe consequences when requesting the parcellation consistency across "neighbors" and confuse the message aggregation during the parcellation model learning.

To address this issue, it is highly necessary to respect the cerebellar geometry, which provides macro-measurable imaging clues characterizing the cytoarchitecture of the cerebellar cortex. Therefore, in this paper, ***first,*** we reconstruct the cerebellar surfaces to restore its convoluted geometry. Specifically, we reconstruct the inner cerebellar surface (the interface between the cerebellar gray matter and white matter), and the pial/outer cerebellar cortical surface (the interface between the cerebellar gray matter and cerebrospinal fluid (CSF)). The reconstructed cerebellar surfaces are represented by triangular meshes, which can be regarded as an undirected graph. ***Second,*** we develop a graph convolutional neural network-based method to do the parcellation on the reconstructed cerebellar surfaces. Specifically, we first extract the informative surface patches from the cerebellar surfaces with the neighborship determined by their intrinsic geodesic distance. And then, we feed these surface patches into the graph convolutional neural network to train a parcellation model, which learns a highly nonlinear mapping from the geometric features of the training patches to the patch labels. One practical challenge for learning this mapping is the need for a large amount of manually labeled data, which is extremely expensive and time-consuming for vertex-level labeling. Therefore, we split the mapping learning into two steps, in the first step, we use massive patches to learn effective low-dimension representations in the latent space. Leveraging the concept of contrastive learning, we require similar representations for the patches from the same region, while distinct representations for patches from different regions. Then, in the second step, we learn the mapping from the low-dimensional latent representation to the parcellation labels. Of note, the representation learning only requires a tag to denote whether the patches are from the same region or not, which is clearly a weaker supervision compared to the accurate patch-level labels.

This paper makes three contributions: 1) we propose a practical pipeline to reconstruct the cerebellar cortical surface that respects its intrinsic geometry; 2) we conduct the parcellation on the original reconstructed cerebellar cortical surfaces, without the

requirement to project them onto a simplified shape, like the sphere, which will inevitably introduce distortions during the projection; 3) we leverage the self-representation learning with weak supervision information to reduce the manual labeling cost for the parcellation network training. To the best of our knowledge, this is the first method that conducts the cerebellar cortex parcellation directly using the original reconstructed cerebellar surfaces.

(a) Inhomogeneity corrected T1w image

(b) Skull stripping and cerebellum extraction

(c) Cerebellum tissue segmentation

(d) Topology correction

Inner surface

Average convexity

Mean curvature

curvedness

Outer surface

Spectral embedding

Shape index

Parcellation map

(e) Cerebellar surface reconstruction

(f) Cerebellar surface geometric features and parcellation map

(g) Cerebellar surface in brain

Fig. 1. Cerebellar surface reconstruction pipeline and the typical geometric features computation.
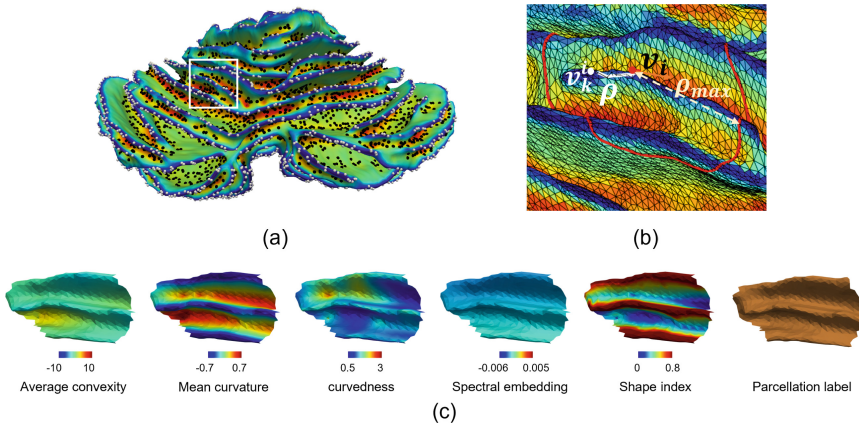
2 Method

Our cerebellar cortex parcellation method can be divided into 3 steps: 1) we reconstruct the cerebellar surfaces and compute the salient geometric features to characterize the geometry of the cerebellar surfaces; 2) we use weakly supervised information to learn effective latent representations from massive informative surface patches. These representations sit in low dimensional latent space, and they are similar for patches from the same region, while are distinct for patches from different regions; 3) we learn a straightforward mapping from the low dimensional latent space to the parcellation labels.

2.1 Cerebellar Surface Reconstruction and Geometric Feature Computation

Given the tissue map of the cerebellum [4], we can reconstruct the cerebellar cortical surfaces, including the inner cortical surface and the pial/outer cortical surface. The pipeline is illustrated in Fig. 1. Specifically, we first do the cerebellum tissue segmentation and then correct the topological errors. Then, we reconstruct the inner cerebellar surface and deform it to the outer cerebellar surface. This framework is similar to the typical cerebral cortical surface reconstruction pipelines. We adopt this typical strategy because, a) it can well preserve the geometry of the highly folded cerebellar cortex; b) it can

establish the vertex-wise correspondence between the inner and outer cerebellar surfaces, which will substantially facilitate the subsequent analysis. More details about widely used cortical surface reconstruction pipelines can be referred to, such as FreeSurfer [5], dHCP [6], FastSurfer [7], HCP [8], and iBEAT V2.0 [9].

With the reconstructed cerebellar surfaces, we can then compute typical geometric features, including the average convexity, mean curvature, curvedness, shape index, and Laplacian spectral embedding [10]. These geometric features encode the local and global geometries of the cerebellar sulci and gyri, which can be used to identify the boundary of each parcellation region (typically appearing at the sulcal bottoms of the cerebellar surface). Specifically, the average convexity encodes the integrated normal movement of a vertex during inflating the inner cerebellar surface and mainly reflects the coarse-scale geometrical information of cerebellar cortical folding [11]; the mean curvature is computed as the average of the minimal and maximal principal curvatures of the inner cerebellar cortical surface and encodes the fine-scale local geometric information of cerebellar cortical folding [11]; also, based on the minimum and maximum principal curvatures, we can compute the curvedness and shape index [12], which characterize the local shape information of the cerebellar cortical surface [11]; the Laplacian spectral embedding reflects the vibration mode of a graph [10], which can be used as a global feature to characterize the cerebellar cortex. Figure 1(f) visualizes these geometric features on the inner cerebellar cortical surface and the corresponding parcellation map. In addition, since the cerebellar cortical surface is not separated into left and right hemispheres, we computed the vertex-wise distance to the geometric center of the surface as additional features, named centroid relation features, which can provide useful clues for the localization of symmetric regions.

(a)

(b)

v_k^i ρ v_i ρ_{max}

-10   10
Average convexity

-0.7   0.7
Mean curvature

0.5   3
curvedness

-0.006   0.005
Spectral embedding

0   0.8
Shape index

Parcellation label

(c)

Fig. 2. Intrinsic surface patch extraction. (a) The patch sampling positions on the cerebellar surface. The white dot indicates the gyri samples, while the black dots indicate the sulci samples. (b) A local intrinsic patch bounded by the geodesic distance ρ_{max}. All vertices whose geodesic distance to the patch center lower than ρ_{max} will be included in the patch. (c) Typical geometric features on the extracted local intrinsic patch in (b).

2.2 Weakly Supervised Cerebellar Patch Representation Learning

Provided with the computed geometric features of the cerebellar surface, we can train a neural network to conduct the parcellation. However, treating the entire cerebellar surface as a single instance would require dense manual labeling for the network training, which is time-consuming and labor-intensive. Therefore, we choose the patch-wise strategy for the parcellation network training. The advantage is that it can significantly enlarge the training samples, while this comes at the cost of requiring the model to possess strong localization ability for the local patch. Directly learning a mapping from the geometric features of a local patch to the parcellation labels is very challenging for the cerebellar surface due to the shape complexity. The local sulci of the cerebellar surface are deeper, and the gyri appear more consistent shape patterns, compared to the cerebral cortical surfaces. This indicates that more training samples are required to enable the network to possess the localization ability from the local patches for a robust parcellation model.
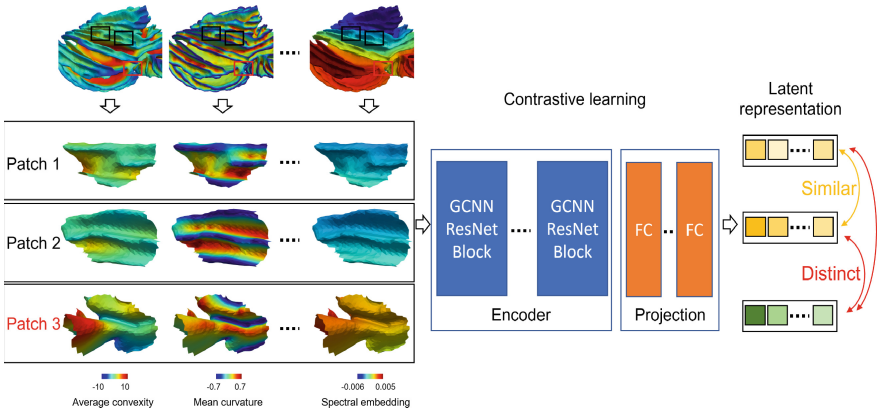
Fig. 3. Weakly supervised representation learning framework. After learning, we enforce similar latent representations between patches 1 and 2 (coming from the same region), but distinct representation between patches 1 (or 2) and patch 3.

To further reduce the cost of the expensive manual labeling, we split the parcellation network training into two steps. In the first step, we learn distinctive latent representations for patches from different regions in a weakly supervised manner. Specifically, we enforce similar representations for patches from the same region, while distinct representations for patches from different regions. This can be achieved with a contrastive learning framework [13]. Herein, we name it weakly supervised since in this step, we only require a tag to denote whether the patches are from the same region or not and we do not require the precise labels of which region it belongs to.

Following the above motivation, we can formulate the weakly supervised patch-based representation learning below. We first extract the local patches from the cerebellar surface, which can be regarded as an undirected graph (V, E), with V being the vertex set and E being the edge set. For any vertex, we can extract a local patch, which contains a set of vertices v_k^i whose geodesic distances to v_i are bounded by the predefined maximal

geodesic distance ρ_{max}, as illustrated in Fig. 2. Considering the prior knowledge that parcellation boundaries of the cerebellar surface generally appear at the sulcal fundi or gyral crests, we sample more patches from those locations, and meanwhile reduce patch samples for vertices sitting on sulcal walls. This can be achieved by thresholding the mean curvature map of the cerebellar surface, since the sulcal fundi and gyral crests typically have larger curvature magnitude, as illustrated in Fig. 2(a), where white dots denote the sample points near gyri crests, and black dots denote the sample points near sulcal fundi. Figure 2(c) shows a typical surface patch with various geometric features and parcellation labels.

Once the local patches are extracted, we feed them into a graph convolution based neural network to learn their latent representation. Herein, we use the widely used residual network [14] as the basic block for constructing our latent representation learning framework. Figure 3 illustrates the learning framework. Specifically, given 3 local patches (p_1, p_2, p_3), where p_1 and p_2 are from the same region and p_3 is from a different region. After the encoding, we can denote their latent representations to be (z_1, z_2, z_3). Then, we enforce (z_1, z_2) to be similar (using the cosine criterion) in the latent space while (z_1, z_3) and (z_2, z_3) to be distinct. This can be achieved by minimizing the normalized temperature scaled cross entropy (NT-Xent) loss, which is defined as:

$$\mathcal{L}(p_1, p_2, p_3) = -log \frac{e^{\cos(z_1, z_2)/\tau}}{e^{\cos(z_1, z_2)/\tau} + e^{\cos(z_1, z_3)/\tau} + e^{\cos(z_2, z_3)/\tau}}$$

where τ is a hyperparameter (named temperature). It is worth noting that, a) it has been validated that this loss has superior performances in many representation learning tasks than other contrastive losses [13]; b) during the training, we can select multiple triplet patches into a single batch and use the average loss over the entire batch to make the training more robust. Since we have extracted a large number of patches from the cerebellar surface, after combining patches according to their tags, we can obtain an even larger set of training samples to ensure network convergence.

2.3　Mapping from Latent Space to Parcellation Labels

Through representation learning, we have encoded the patches with multiple channels of geometric features into the latent space. This step *not only* makes the extracted patches more distinct in the latent space from the regional perspective *but also* significantly reduces the potential feature dimension, which can greatly facilitate the parcellation network training, compared to training the parcellation network directly from the patch-wise geometric features.

Therefore, given the patch-wise latent representations, we further train a multilayer perceptron (MLP) to accomplish the parcellation task. Specifically, we use the parcellation labels from the manually labeled cerebellar surface to supervise the MLP training. Herein, the patch center label is used as the patch label to train the MLP and the popular cross-entropy loss is adopted.

However, since each patch is parcellated independently without considering the spatial consistency, it is possible to generate isolated parcellation labels and cause inconsistency in a geodesic neighborhood. To improve the parcellation, we further use the

graph cuts method [15] to explicitly impose spatial consistency. Specifically, since most of the regions are separated at the sulci fundi of the cerebellar cortex according to the manual labeling protocol, we explicitly formulate parcellation as a cost minimization problem, i.e., $E = E_d + \lambda E_s$. Here, E_d is the data fitting term, which is defined as: $E_d = -\log p_v(l_v)$, where $p_v(l_v)$ is the probability of assigning vertex v as label l_v; E_s is the smoothness term, which is defined as: $E_s = \sum_{v^* \in \mathcal{N}_v} C_{v,v^*}(l_v, l_{v^*})$, where vertex v^* is the direct neighbor of v, and $C_{v,v^*}(l_v, l_{v^*})$ is the cost to label vertex v as l_v and also label vertex v^* as l_{v^*}. Herein, we used the formula from [16] to define $C_{v,v^*}(l_v, l_{v^*})$; finally, λ is a weight used to balance them.

Table 1. Performance comparison by different methods.

Comparison methods	Dice ratio
GCN [20]	76.03 ± 3.20
GIN [21]	75.20 ± 2.40
Graph U-Net [22]	73.17 ± 4.46
Graph SAGE [23]	73.52 ± 4.69
Proposed	81.26 ± 2.37

3 Experiments

3.1 Dataset and Implementation

To validate our method, we manually labeled 10 subjects from BCP dataset [17] in a vertex-wise manner using an inhouse developed toolkit. Specifically, for each cerebellum MRI, we first segmented it into white matter, gray matter, and CSF using a deep convolutional neural network [4]. Then, following the processing pipeline in Sect. 2.1, we reconstructed the geometrically accurate and topologically correct cerebellar surfaces [18, 19]. Finally, each inner cerebellar surface is labeled into 17 lobules following the SUIT parcellation strategy [3] by an expert. To ensure the labeling quality, each manually labeled region is further validated and cross-modified by another 2 experts to alleviate the subjective bias and ensure the labeling quality. Due to the limited subject number, we adopted these manually labeled cerebellums to validate our method in a 10-fold cross validation.

We implemented our method mainly based on the pytorch (https://pytorch.org/) and pytorch geometric (https://pyg.org/) packages. The graph convolution operator [20] implemented in the pytorch geometric package is adopted as the major building block to construct our network. For the contrastive learning framework, we used 18 ResNet blocks for the encoding and 3 fully connected layers for the projection. After the cerebellar cortical surface reconstruction, each scan typically has around 90k vertices. We extract 8k patches from each training scan. The ρ_{max} is set to 15mm. For the NT-Xent loss, the hyperparameter τ is set to 0.5. For the graph-cut cost, the weight balance λ is set to 1. For the testing subject, we perform the parcellation for each patch.

3.2 Comparison with the State-of-the-Art Methods

Since we conducted our parcellation on the original cerebellar surface, which can be regarded as a graph, we compared our method with state-of-the-art graph convolutional neural network based methods [20–23] for the cerebellum parcellation task. The Dice similarity coefficient (DSC) between the predicted label and the manual label is adopted as a quantitative evaluation of the parcellation performance. Table 1 reported the average DSC of the parcellation for all regions acquired by our proposed method and the comparison methods. It can be seen that our method achieves best performance, indicating the effectiveness of the latent patch feature embedding obtained by our proposed weakly supervised contrastive learning strategy.

Table 2. Different features' influence on parcellation performance.

Comparison methods	DSC
Raw geometric features	46.02 ± 4.13
Raw geometric features + spectral feature	72.72 ± 4.52
Raw geometric features + spectral feature + centroid relation feature	81.26 ± 2.37

3.3 Different Features' Influence Analysis

We also validated the influence of each geometric feature on the final parcellation performance. We conducted the ablation study by removing parts of the geometric features for the parcellation. Table 2 reported the average DSC using different feature combinations. It can be seen that the raw geometric features, i.e., the average convexity, mean curvature, curvedness, and shape index, are difficult to capture the localization ability. This is reasonable since the cerebellar cortex has much deeper sulci, and appear more similar geometric patterns, compared to the cerebral cortex. Therefore, the local patches have relatively low localization ability. However, after adding the spectral feature and the centroid relation feature, the parcellation performance is greatly improved.

4 Conclusion

In this paper, we propose an automated method for anatomically meaningful cerebellar cortical surface parcellation. We firstly reconstruct the geometric accurate and topologically correct cerebellar surfaces and then compute several widely used geometric features to comprehensively characterize the geometries of the cerebellar surface. Next, we extract local surface patches from the reconstructed cerebellar surfaces with the neighborship defined by the intrinsic geodesic metric. These extracted local surface patches are projected to a low dimensional latent space with a contrastive learning framework, i.e., patches from the same region are enforced to have similar representations, while patches from different regions have distinct representations. After that, we train a neural

network to map the latent representations to the parcellation labels. Comparison to the state-of-the-art methods has validated the superior performance of our method. Currently, our work has two limitations, a) the quantitative evaluation is based on a small number of the subjects, due to the expensive manual labeling cost. In the future, we plan to involve and release more manually labeled cerebellar cortical surfaces to further improve the generalizability of the current framework, enhance the representation learning and validate our method on larger datasets; b) a graph cut post-processing is needed to remove potential inconsistent labelling. In the future, we plan to directly add the neighborhood smooth labeling constraint into the network cost function to obtain an end-to-end cerebellar cortical surface parcellation method.

References

1. Davie, C.A., et al.: Persistent functional deficit in multiple sclerosis and autosomal dominant cerebellar ataxia is associated with axon loss. Brain **118**, 1583–1592 (1995)
2. Riva, D., Giorgi, C.: The cerebellum contributes to higher functions during development evidence from a series of children surgically treated for posterior fossa tumours. Brain **123**, 1051–1061 (2000)
3. Carass, A., et al.: Comparing fully automated state-of-the-art cerebellum parcellation from magnetic resonance images. Neuroimage **183**, 150–172 (2018)
4. Sun, Y., Gao, K., Niu, S., Lin, W., Li, G., Wang, L.: Semi-supervised transfer learning for infant cerebellum tissue segmentation. In: Liu, M., Yan, P., Lian, C., Cao, X. (eds.) MLMI 2020. LNCS, vol. 12436, pp. 663–673. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59861-7_67
5. Fischl, B.: FreeSurfer (2012)
6. Makropoulos, A., et al.: The developing human connectome project: a minimal processing pipeline for neonatal cortical surface reconstruction. Neuroimage **173**, 88–112 (2018)
7. Henschel, L., Conjeti, S., Estrada, S., Diers, K., Fischl, B., Reuter, M.: FastSurfer - a fast and accurate deep learning based neuroimaging pipeline. Neuroimage **219**, 117012 (2020)
8. Glasser, M.F., et al.: The minimal preprocessing pipelines for the human connectome project. Neuroimage **80**, 105–124 (2013)
9. 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. 2023. 1–32 (2023)
10. Lombaert, H., Sporring, J., Siddiqi, K.: Diffeomorphic spectral matching of cortical surfaces. In: Gee, J.C., Joshi, S., Pohl, K.M., Wells, W.M., Zöllei, L. (eds.) IPMI 2013. LNCS, vol. 7917, pp. 376–389. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38868-2_32
11. Fischl, B., Sereno, M.I., Dale, A.M.: Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. Neuroimage **9**, 195–207 (1999)
12. Goldman, R.: Curvature formulas for implicit curves and surfaces. In: Computer Aided Geometric Design, pp. 632–658. North-Holland (2005)
13. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.: A simple framework for contrastive learning of visual representations. In: International Conference on Machine Learning, pp. 1575–1585. International Machine Learning Society (IMLS) (2020)

14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
15. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Trans. Pattern Anal. Mach. Intell. **26**, 1124–1137 (2004)
16. Wu, Z., et al.: Registration-free infant cortical surface parcellation using deep convolutional neural networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 672–680 (2018)
17. 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)
18. Sun, L., et al.: Topological correction of infant white matter surfaces using anatomically constrained convolutional neural network. Neuroimage **198**, 114–124 (2019)
19. Li, G., et al.: Measuring the dynamic longitudinal cortex development in infants by reconstruction of temporally consistent cortical surfaces. Neuroimage **90**, 266–279 (2014)
20. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: International Conference on Learning Representations. International Conference on Learning Representations. ICLR (2017)
21. Xu, K., Jegelka, S., Hu, W., Leskovec, J.: How powerful are graph neural networks? In: 7th International Conference on Learning Representations, ICLR 2019. International Conference on Learning Representations. ICLR (2019)
22. Gao, H., Ji, S.: Graph U-nets. IEEE Trans. Pattern Anal. Mach. Intell. **44**, 4948–4960 (2022)
23. Hamilton, W.L., Ying, R., Leskovec, J.: Inductive representation learning on large graphs. In: Advances in Neural Information Processing Systems, pp. 1025–1035. Neural Information Processing Systems Foundation (2017)