International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
Impact Factor (2020): 6.733

Segmentation of Cerebral Cortex Based on 3D CNN

Peng Liu[1]

[1]North China Electric Power University, School of Control and Computer Engineering, No.2 Beinong Road, Changping District, Beijing, China
15210266207[at]163.com

Abstract: ***Research on the surface of the cerebral cortex is crucial to understanding how the brain works. In this study, a three-dimensional full convolution neural network (3D CNN) based on sample equalization mechanism is proposed for MRI cortical region segmentation. We implement a deeper network structure through a small kernel, and obtain multi-scale information by integrating local and global context information to improve the segmentation accuracy. The algorithm is verified by comprehensive experiments on Mindboggle dataset. The results show that compared with other recent methods, this method improves the segmentation accuracy.***

Keywords: Deep Learning, Cerebral Cortex Segmentation, 3D CNN

1. Introduction

When neural image analysis can study the function of brain geometric structure and anatomical information. Magnetic resonance imaging provides obvious cortical geometry information, which plays a very important role in human cognition, vision and perception. When studying all aspects of the brain, the study of the surface of the cerebral cortex is of great significance. The accurate segmentation of the structure of the cerebral cortex is very important for the study of various brain diseases and the evaluation of brain structural abnormalities [1]. So far, some effective neuroimaging analysis tools, such as FreeSurfur [2], SPM [3], ants [4] can make fine processing of cerebral cortex data. These tools usually use multiple image transformation steps, some of which require careful fine-tuning of parameters, such as convergence threshold, smoothing level or number of iterations [5]. In addition, due to extensive numerical optimization, such as nonlinear registration, these methods have high computational cost and long running time. It takes several hours to process a single image, which greatly limits large-scale medical research. Automatic segmentation of cerebral cortex is a great challenge. Clinicians mainly rely on manual marking, which is a very time-consuming process, requires professional doctors to mark, and is prone to inconsistency [6]. Therefore, there is an urgent need for a fast, accurate, repeatable and fully automated method to segment the structure of cerebral cortex.

Deep learning method provides a new way to solve cerebral cortex segmentation [7], and has great speed advantage. Different from the classical model based on Feature Engineering, the deep learning model directly learns complex features from data, and extracts features through trainable convolution kernel and optional pool operation. For example, fully convolutional neural network (FCNN) [8] can learn the correct feature representation from the image itself in an end-to-end manner without complex preprocessing steps. These methods can be parallelized effectively on the graphics processing unit (GPU), resulting in a huge speed increase. The task of whole brain segmentation is particularly challenging due to the spatial dependence between complex 3D structures and slices, the large number of labels, the limitation of memory requirements and the variability of subjects cerebral cortex structure.

At present, there are mainly several methods based on deep learning for specific tasks, such as tumor segmentation [9] [10] [11], brain injury segmentation [12] [13], image reconstruction [14] [15] [16], brain disease prediction [17] [18], etc. But so far, only a few teams have achieved more than 25 categories of cerebral cortex segmentation [19] [20] [21]. Most brain segmentation networks are trained on extracted 3D slices or 2D slices [7], but these methods will more or less lose the spatial information that is crucial to the classification results. However, due to the limitation of memory resources, it is impossible to construct a full 3D deep neural segmentation network trained with whole brain MRI and with a large number of labels. Another atlas based method is to align one or more anatomical templates with the target image through linear or nonlinear registration process, and then transfer the segmentation label from the template to the image []. Although these methods usually provide satisfactory results, the segmentation time is usually long (from minutes to hours), which is due to the complexity of the registration step. The methods based on point cloud coordinate segmentation usually transform the 3D brain data extraction into point coordinate form. These methods will save a lot of time, space and memory requirements. However, these methods completely ignore the spatial information in the transformation process, so the final result often loses a lot of context spatial information, which will directly affect the experimental results.

Kamnitsas et al, proposed a full 3D method [12], which consists of a 3D CNN that generates soft segmentation results and a fully connected 3D CRF that applies generalization constraints and obtains the final label. The idea of kamnitsas [12] is adopted here to realize a deeper architecture by using a small kernel. In addition, the connection between local and global contexts is established by using a variety of multi-scale methods. By embedding the middle layer output in the final prediction, this method maintains the consistency between the extracted features at different scales, and directly embeds the image information of different resolutions in the segmentation process []. Different from the previous network structure, the global context is modeled using a separate path and low-resolution image. Due to the need for detailed segmentation of the whole cerebral cortex structure, more global and local context information can be collected by fusing multiple resolution information, which is of great significance for the hint of cerebral cortex boundary region segmentation.

In order to verify the effectiveness of our net for cerebral cortex segmentation results, a comprehensive experiment was carried out on the public mindboggle [24] data set to illustrate the learning ability of the experimental method. Mindboggle [24] dataset provides the established benchmark and publicly available manual label dataset. The segmentation results obtained by our method are highly consistent with the segmentation results of standard manual label dataset.

2. Method

This section describes the details of our net. Firstly, we describes the basic structure of the network (3D FCNN), and obtains a deeper network structure combined with the idea of small kernel, so that the network can learn more complex feature hierarchy and reduce the risk of over fitting. The next part focuses on the use of multi-scale strategy to obtain multi-scale information from the middle convolution layer to improve this architecture and improve the training accuracy. The overall structure of the network is shown in Figure 1.

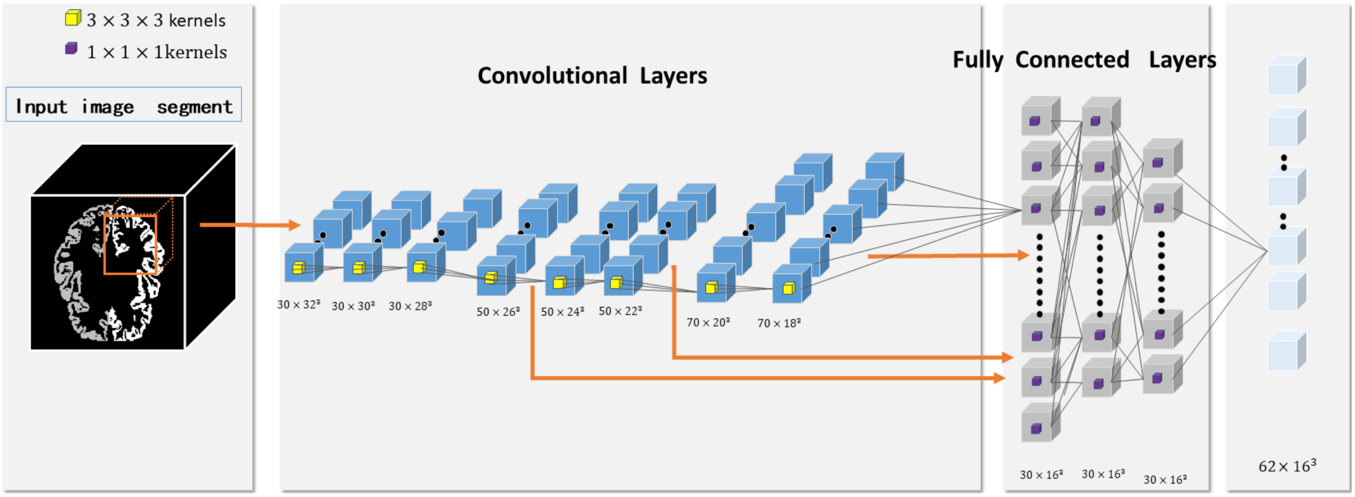

Figure 1: Multi-scale 3D CNN based on multi-resolution feature fusion

2.1 3D FCNN network with small kernel

Network depth is the key parameter of the system, and the deeper network has stronger recognition ability. Although deep networks may be more difficult to train than shallow networks [25], using more layers can improve performance. One of the limitations of 3D architecture is the harsh requirements on memory [26]. When deep network is used, pooling layer is usually used to reduce the size of middle layer to solve memory constraints. Another key parameter limited by memory requirements is the number of filters per layer, especially in the network layer with higher dimensions. Finally, the size of input and batch processing also needs to be suitable for hardware memory. Another limitation of 3D networks is that 3D convolution is computationally expensive and will multiply the number of parameters. This problem can be alleviated by using smaller filters in each convolution layer.

In order to construct a deeper 3D network structure, we use a small convolution kernel of 3^3, which contains fewer weights. This design method was previously found to be beneficial to the classification of natural images [27], and its effect on 3D networks is more significant. If the convolution kernel of 5^3 is used in the basic 3D FCNN network, the network parameter is $5^3/3^3$=4.63 times that of 3^3 convolution kernel. Therefore, by using a smaller convolution kernel to replace each layer of the basic FCNN network, a deeper network structure can be designed.

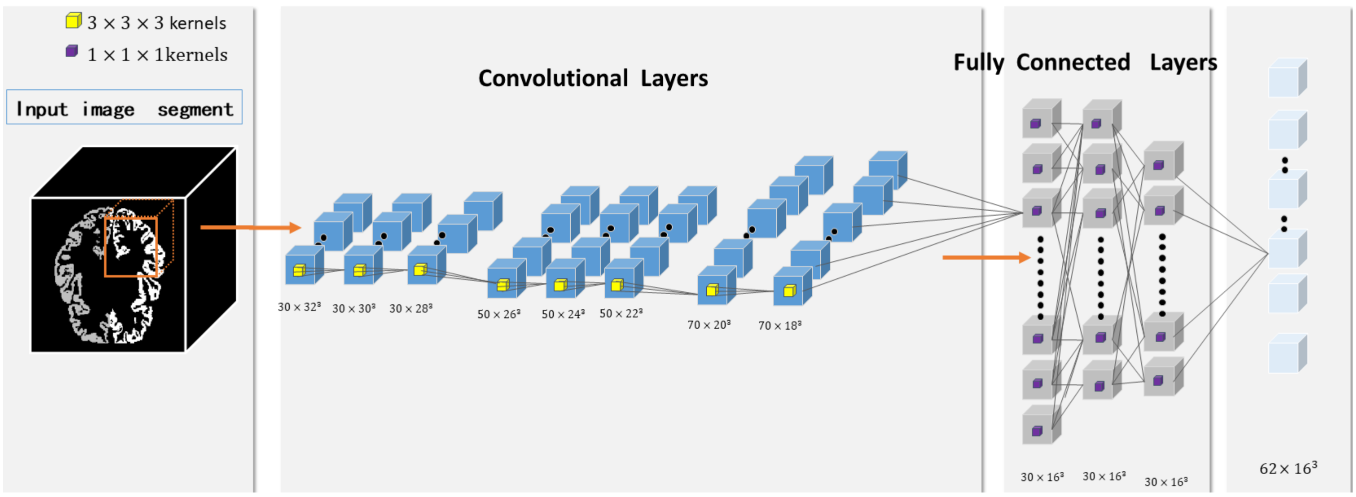

Figure 2: 3D FCNN network structure with small kernel

The segmentation method of 3D full CNN (FCNN) architecture with small kernel is shown in Figure 2. The architecture consists of 8 convolution layers, each layer contains several 3D convolution cores. Use m_l to represent the convolution kernel number of each layer of the network, then the output of each layer is:

$$y_l^k = f(\sum_{n=1}^{m_{l-1}} W_i^{k,n} \otimes x_{l-1}^n + b_l^k) \quad (1)$$

The feature maps generated by convolution in each layer are slightly smaller than their input volume. If the convolution kernel size of each layer is 3^3, the size difference in each dimension is equal to the convolution kernel size in this dimension minus 2. The network adopts prelu activation function, which is defined as follows:

$$f(x_i) = \max(a_i x_i, x_i) \quad (2)$$

In the standard CNN, the full connection layer is added to the end of the network to encode semantic information. However, in order to ensure that the network contains only the convolution layer, the method described by deepmedic [12] is used here, in which the full connection layer is converted into a large 1^3 convolution set, which allows the network to retain spatial information and learn the parameters of these layers, just as in other convolution layers. The activation of neurons in the last layer corresponds to a specific segmentation class label. Finally, it is transformed into normalized probability value by softmax function, and the final segmentation result is obtained. The probability of each category is as follows:

$$p_a = \frac{\exp(Z_L^a)}{\sum_{a_1=1}^{a} \exp(Z_L^{a_1})} \quad (3)$$

2.2 Cross multiscale fusion features

In CNN, the sequence of layers encodes features representing an increasing level of abstraction; the first convolution layer usually simulates simple edge information, while the deeper convolution layer before the full connection layer directly abstracts larger and more complex structures. Combining multi-scale features has been found to be beneficial in other recent studies [28] [29] in which the whole image is processed in the network, a small amount of convolution is applied, and then the feature images of different sizes are unified into the same proportion by down sampling. According to this principle, we further improve the 3D FCNN to include multi-scale information in the segmentation results.

The multi-scale version of our segmentation model is shown in Figure 1. As shown in the net, there is a single 3D image as the input. Here, we use the highest resolution MRI image and integrate the feature map from the middle layer of partial convolution layer into the full connection layer. Compared with other multi-scale methods, this strategy has two important advantages, that is the input image is resampled at multiple scale resolutions before being input into the network. First, because it has a set of convolution filters in each layer, rather than one in each path, the features on different scales are more likely to be consistent. In addition, because the features of the middle layer are injected into the top layer, fine-grained information is directly used in the segmentation process.

Because each feature map comes from different middle layers, it is necessary to unify the feature maps of different sizes at the end of the network, and the middle feature map needs to be down sampled to match the size of the feature map of the last layer. These feature maps are compatible in size and represent different resolutions respectively. In addition, their input image receptive fields are also different (that is, the middle layer feature map has a smaller receptive field and the last layer convolution layer has a larger receptive field. In this way, the local information and context information are effectively combined to greatly improve the segmentation results.

3. Experiment

In this section, a series of experiments are mainly introduced to analyze the influence of experimental methods, and the experimental parameters finally adopted in our network are introduced. It focuses on the influence of multi-scale training on our experimental results.

3.1 Experimental data and environment

We now validate our multiscale learning method. We benchmark our experimental performance using Mindboggle, a publicly available artificially labeled cerebral cortex data set. It included 101 subjects collected from different websites. We used 3D MRI images to divide the left and right brain into 31 regions according to the dkt31 protocol. The experiment was conducted on an i7 desktop with 16GB ram and NVIDIA geforce GTX 1660ti GPU.

Mindboggle-101 is brain MRI data sets compiled by Arno et al. in 2012.101 3D brain MRI images were integrated from the existing public data set. All images were taken from healthy volunteers and from several independent data sets. The vertices of cortical surface were assigned a given label based on local surface curvature and average convexity, a priori label probability and adjacent vertex labels. In this paper, a total of 80 images from four data sets are selected for the experiment. The specific conditions are shown in Table 1:

Table 1: Experimental data sheet

Dataset name	Number of images	Subject age (mean)
MMRR-21	21	22-61 (31.8)
OASIS-TRT-20	20	19-34 (23.4)
NKI-RS-22	22	20-40 (26.0)
NKI-TRT-20	20	19-60 (31.4)

The image and label image are shown in Figure 4. The label data is the label of cerebral cortex region, which is marked by professionals according to DKT protocol. The left and right half regions of the brain in the image contain 31 cortical regions respectively, a total of 62 cortical regions.

Figure 3: MRI image and label image of Mindboggle-101

Since our experiment here only focuses on the thin outer cortex of the brain, the relevant information of other brain data is not our concern, such as gray matter areas, which need to be removed. In addition, there are great differences in the size and shape of the original data, so before the experiment, we need to accurately preprocess the experimental data, which is directly related to the accuracy of our experimental results. The processing flow includes image clipping, label continuity, gray normalization, ROI region image construction and so on.

3.2 Comparison of experimental results

The final structure of our network is obtained by fusing the middle layer information in the last layer of each type of network. Three hidden layers are added before the classification layer to combine the multi-scale features to obtain a network with four types of networks, and the deepest degree is 11 layers (see Figure 1). In the first type of network, the final feature map is obtained by integrating the network outputs of layers 4, 6 and 8 as the input of the full connection layer. The feature map from layers 4 and 6 needs to be down sampled to match the size of the feature map of layer 8 network.

Our first experiment focused on comparing the contribution of multi-resolution features to the final cerebral cortex segmentation. We expand the number of network layers of 3D FCNN network to make it have the same number of network layers as the network structure in the figure above. By deleting the lower layer prediction (i.e. skip-connection), only the core network (the network in Fig.1) is left, which we name DeepCNN. The following figure shows the comparison of experimental results without adding multi-scale information and adding multi-scale information. In order to confirm that our experimental results are improved due to the adoption of multi-scale idea, we conduct experiments by doubling the feature map of network layering, increasing the depth of network layer (deep++) and fusing more middle layer information (skip++). It can be found that the final experimental results have not improved, and even there are signs of fitting. Moreover, the average dice score decreases when more middle layer information is fused, and the experimental results are shown in Table 2.

Table 2: Comparison of experimental results

	Avg. Dice (%)	Avg. Hausdorff (mm)
DeepCNN	0.8264	1.97
FreeSurfer	0.8439	2.11
Our	**0.8508**	**1.82**
Feature map*2	0.8495	1.87
Deep++	0.8489	1.88
Skip++	0.8312	2.26

Figure 4 shows the segmentation results obtained by using our net, in which different colors represent different types of blocks. The left figure is the segmentation results, and the figure is the ground truth comparison diagram. It can be seen from the figure that the segmentation results obtained by our method are highly similar to the ground truth, and the segmentation smoothness effect is very good.

Figure 4: Segmentation results of cerebral cortex

3.3 Adding residual error and Sample Equalization Mechanism

As the depth of the network deepens, it brings another problem, that is, the degradation of network performance. When the depth deepens, the error rate increases. Here, we use the residual proposed by Resnet [30] to design and solve the degradation problem. At the same time, it also solves the gradient problem and improves the performance of the network. Here, the input of layer 3, 5 and 7 of the network together with its output are used as the input of the next layer to ensure the performance of the network.

In addition, our cortical segmentation task has many categories, and the number of each category varies greatly. The category imbalance data may lead to better segmentation results for the most common category in the final segmentation results, while the category with less sample points has poor segmentation results. Therefore, the method of counting the number of samples is introduced into the loss function, which is named lcc loss to ensure that the points with a small number of samples are trained. Our loss function consists of cross entropy loss function and lcc loss, which are defined as follows:

$$L(G,Y) = 1 - \frac{1}{I}\sum_{i=1}^{I}\sum_{j=1}^{J} G_{ij}\log Y_{ij} - \text{similarity}(G,Y) \quad (4)$$

$$\text{similarity}(G,Y) = \frac{\sum_{i=1}^{n}\left(\sum_{j=1}^{J} G_{ij} \sum_{j=1}^{J} Y_{ij}\right)}{\sum_{i=1}^{I}\left(\sum_{j=1}^{J} G_{ij}\right)^{2} \sum_{i=1}^{I}\left(\sum_{j=1}^{J} Y_{ij}\right)^{2}} \quad (5)$$

Where I is the number of samples, J is the number of classifications, and G_{ij} is the one hot code of the ground truth of point I, Y_{ij} is the output probability of training. The comparison with the previous experimental results is shown in Table 3:

Table 3: Comparison of experimental results

	Avg. Dice (%)	Avg. Hausdorff (mm)
DeepCNN	0.8264	1.97
FreeSurfer	0.8439	2.11
Our	**0.8508**	**1.82**
Our + lcc loss	**0.8524**	**1.78**

4. Conclusion

We present an automatic segmentation method of cerebral cortex based on full convolution network. The training scheme not only has high calculation efficiency, but also ensures the training accuracy. By using a small convolution kernel, a deeper network with fewer parameters can be obtained, and it is less prone to over fitting. In addition, by injecting the output of the middle layer into the full connection layer of the network, the local and global context information is obtained, and the features extracted at different scales have consistency, and the local information is embedded in the results to obtain good segmentation results.

By describing the network idea and verifying it with the public human brain MRI data set, the experimental evaluation shows that this method has a good effect on the segmentation of cerebral cortex.

References

[1] P. Koolschijn, N. E. van Haren, G. J. Lensvelt-Mulders, H. Pol, E. Hilleke, R. S. Kahn, Brain volume abnormalities in major depressive disorder: Ameta-analysis of magnetic resonance imaging studies, Human brain mapping 30 (11) (2009) 3719–3735
[2] Fischl, B., 2012. Freesurfer. Neuroimage 62 (2), 774–781.
[3] Friston, K., Ashburner, J., Kiebel, S., Nichols, T., Penny, W., 2007. Statistical Parametric Mapping: the Analysis of Functional Brain Images. Academic Press.
[4] Avants, B. B., Tustison, N., Song, G., 2009. Advanced normalization tools (ants). Insight Journal 2, 1–35.
[5] Henschel L, Conjeti S, Estrada S, Diers K, Fischl B, Reuter M, FastSurfer-A fast and accurate deep learning based neuroimaging pipeline, NeuroImage 219 (2020), 117012.
[6] M. Deeley, A. Chen, R. Datteri, J. Noble, A. Cmelak, E. Donnelly, A. Malcolm, L. Moretti, J. Jaboin, K. Niermann, et al., Comparison of manual and automatic segmentation methods for brain structures in the presence of space-occupying lesions: a multi-expert study, Physics in medicine and biology 56 (14) (2011) 4557.
[7] W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, D. Shen, Deep convolutional neural networks for multi-modality isointense infant brain image segmentation, NeuroImage 108 (2015) 214–224.
[8] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp.3431–3440
[9] Dong, H., Yang, G., Liu, F., Mo, Y., Guo, Y., July. Automatic brain tumor detection and segmentation using u-net based fully convolutional networks. In: Medical Image Understanding and Analysis-21st Annual Conference. MIUA2017, Edinburgh, UK, pp.506-517. Proceedings, 2017
[10] Brainlesion Glioma, 2018. Multiple Sclerosis, Stroke and Traumatic Brain Injuries-4th International Workshop. BrainLes. Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part I, Vol.11383 of Lecture Notes in Computer Science, Springer, 2019
[11] Amin, J., Sharif, M., Yasmin, M., Fernandes, S. L., 2018. Big data analysis for brain tumor detection: deep convolutional neural networks. Future Generat.comput. Syst.87, 290–297.
[12] Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., Glocker, B., 2017. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. Med. Image Anal.36, 61–78.
[13] Chen, X., Konukoglu, E., 2018. Unsupervised detection of lesions in brain MRI using constrained adversarial auto-encoders. In: Proceedings of the 1st International Conference on Medical Imaging with Deep Learning MIDL 2018. Amsterdam, Netherlands, 4-6 July 2018
[14] Mardani, M., Gong, E., Cheng, J. Y., Vasanawala, S., Zaharchuk, G., Alley, M. T., Thakur, N., Han, S., Dally, W. J., Pauly, J. M., Xing, L., 2019. Deep generative adversarial networks for compressed sensing automates MRI. IEEE Trans. Med. Imag.38, 167–179.
[15] Dedmari, M. A., Conjeti, S., Estrada, S., Ehses, P., Stocker, T., Reuter, M., 2018.complex fully convolutional neural networks for mr image reconstruction. In: Knoll, F., Maier, A., Rueckert, D. (Eds.), Machine Learning for Medical Image Reconstruction. Springer International Publishing, pp.30–38
[16] Yang, G., Yu, S., Dong, H., Slabaugh, G. G., Dragotti, P. L., Ye, X., Liu, F., Arridge, S. R., Keegan, J., Guo, Y., Firmin, D. N., 2018. DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE Trans. Med. Imag.37 (6), 1310–1321.
[17] Payan, A., Montana, G., 2015. Predicting Alzheimer's disease-a neuroimaging study with 3d convolutional neural networks. In: ICPRAM 2015-Proceedings of the International Conference on Pattern Recognition Applications and Methods, ume 2. Portugal, Lisbon, pp.355–362, 10-12 January, 2015.
[18] Hosseini-Asl, E., Keynton, R., El-Baz, A., 2016. Alzheimer's disease diagnostics by adaptation of 3d convolutional network. In: 2016 IEEE International Conference on Image Processing. ICIP, pp.126–130.
[19] Roy, A. G., Conjeti, S., Navab, N., Wachinger, C., Initiative, A. D. N., et al., 2019. Quicknat: a fully convolutional network for quick and accurate segmentation of neuroanatomy. Neuroimage 186, 713–727.
[20] Huo, Y., Xu, Z., Xiong, Y., Aboud, K., Parvathaneni, P., Bao, S., Bermudez, C., Resnick, S. M., Cutting, L. E., Landman, B. A., 2019.3d whole brain segmentation using spatially localized atlas network tiles. Neuroimage 194, 105–119

[21] Wachinger, C., Reuter, M., Klein, T., 2018. Deepnat: deep convolutional neural network for segmenting neuroanatomy. Neuroimage 170, 434–445
[22] Karthik Gopinath, Christian Desrosiers, Herve Lombaert, Graph Convolutions on Spectral Embeddings for Cortical Surface Parcellation, Medical Image Analysis, Volume54, 2019, Pages297-305, ISSN1361-8415, https://doi.org/10.1016/j. media.2019.03.012.
[23] Dolz, Jose, C. Desrosiers, and I. B. Ayed. "3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. " Neuroimage (2017): S1053811917303324.
[24] Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., H¨ame, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Chaibub Neto, E., Keshavan, A., 2017. Mindboggling morphometry of human brains. PLOS Computational Biology 13, e1005350
[25] Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks. In: International conference on artificial intelligence and statistics, pp.249–256.
[26] Choromanska, A., Henaff, M., Mathieu, M., Arous, G. B., LeCun, Y., 2015. The loss surfaces of multilayer networks. AISTATS
[27] Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.15
[28] Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.3431–3440.
[29] Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Springer, pp.234–241.
[30] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE conference on computer vision and pattern recognition.2016: 770-778.

Author Profile

Peng Liu, Born in April 1997 in Datong, Shanxi Province, he is a master's student in the school of control and computer engineering of North China Electric Power University. The main research direction is machine learning and medical image processing.