Semantic Segmentation of Cerebral Cortex Based On Pointnet

Yanting Sang[1]

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

Abstract: ***The cerebral cortex is the highest level center for regulating and controlling body movement. The semantic segmentation of the cerebral cortex is of great significance for the study of the neural structure of the brain and the early diagnosis and treatment of brain diseases. At present, most of the existing cerebral cortex segmentation methods are based on the brain structure segmentation based on MRI images, and then extract the cortical information as the final segmentation result. This method has problems such as large amount of data and low segmentation efficiency, and the accuracy is easily affected by the brain. Influence of internal structure segmentation results. In response to this problem, this paper directly takes the cerebral cortex data as the research object, and on the basis of fully considering the characteristics of the cerebral cortex point cloud collection data, the PointNet segmentation network is used to perform semantic segmentation of the cerebral cortex. At the same time, in view of the imbalance of data categories in the dataset, a sample balance mechanism is introduced to improve the loss function of the segmentation model, thereby improving the contribution of smaller categories in the dataset to the segmentation results. The experimental results show that the PointNet network can well solve the problem of semantic segmentation of the cerebral cortex, and the use of the sample equalization mechanism can further improve the accuracy of small category partitioning and improve the segmentation accuracy.***

Keywords: Cerebral cortex, semantic segmentation, PointNet, sample equalization

1. Introduction

The cerebral cortex is composed of multiple semantic divisions with different functions. It is the highest level center that regulates and controls the movement of the body, and plays a decisive role in the life activities of the human body. Precise semantic segmentation of the cerebral cortex helps to better understand the neural structure of the brain, and has important medical value for the early diagnosis and treatment of brain diseases such as autism and Alzheimer's disease.

At present, most of the cerebral cortex segmentation methods are based on brain structure segmentation of MRI images, and then extract cortical information as the result of semantic segmentation segmentation. Common brain structure segmentation methods mainly include threshold-based segmentation methods [1, 2], deep learning-based algorithms [3, 4, 5, 6] and hybrid algorithm-based segmentation methods [7, 8]. The methods to achieve semantic segmentation of cerebral cortex mainly include the method of whole-brain segmentation based on deep learning [5, 6]. However, there are many problems in obtaining cerebral cortex segmentation through brain structure segmentation. First, the data volume of the complete brain structure is large, and the segmentation efficiency is not high; cause interference.

In view of the above problems, it is a beneficial attempt to directly perform semantic segmentation on cerebral cortex data, which can not only reduce the scale of data, but also reduce the influence of other information. The cerebral cortex mainly has two data formats: point cloud collection and patch grid, in which the network data includes topological connection relationship in addition to point collection. In order to better adapt to the two kinds of data, this paper takes the cerebral cortex represented by the point cloud set as the research object, ignoring the topology information in the patch grid.

However, there are few researches on point cloud collections in the medical field. Yang et al. [9] used the 3D intracranial aneurysm dataset Intra to reconstruct point cloud data, and compared the segmentation performance among 11 popular point cloud deep learning networks. In experiments, Gutiérrez et al. [10] used multiple brain structure point clouds for brain structure classification to enable prediction of Alzheimer's disease and mild cognitive impairment. Gazvinia et al. [11] designed a method for tooth semantic segmentation based on the PointCNN model, and achieved high segmentation results. These models have strong pertinence, serve specific human body parts or organs, and are less versatile. On the other hand, the PointNet [12] network has been widely used in indoor scene segmentation [13], large-scale outdoor scene segmentation [14], human body segmentation [15], and hand segmentation [16] due to its ability to directly process unordered point cloud sets and other fields of segmentation. On the basis of fully considering the characteristics of the cerebral cortex point cloud collection data, this paper applies the PointNet network to the cerebral cortex semantic segmentation.

During the segmentation process, due to the uneven distribution of the number of point clouds in different semantic regions of the cerebral cortex, the smaller regions are particularly susceptible to the impact of insufficient sample size, and the segmentation accuracy of this region is difficult to guarantee. But smaller divisions do not imply less importance in medical diagnosis and treatment. Therefore, the sample equalization mechanism OHEM [17], Focal Loss [18], GHM Loss [19] and other methods that have been successfully applied in the field of target recognition are used to improve. The OHEM method only selects 1/3 of the difficult samples to calculate the loss function during training, and the weights of other difficult samples are set to 0, while the GHM loss method defines the extremely difficult samples as outliers during training. Sample and discard it. The processing methods of the above two methods are too simple and crude, and do not meet the characteristic requirements of cerebral cortex data. Under comprehensive consideration, we introduce the Focal Loss proposed by Li et al. [18], reshape the loss function into a lightweight simple sample, and integrate it into the problem of semantic segmentation of the cerebral cortex, so that the training focus is on small partition samples On the premise of not affecting the overall accuracy, the segmentation effect of smaller partitions can be improved to meet the actual needs of medicine.

2. Method

2.1 PointNet

The convolutional network structure in order to achieve weight sharing and other kernel function optimization usually requires a highly regularized format of the input data, and most studies in order to apply irregular point cloud data usually convert it into an ordered 3D voxel grid or image set , but these transformations increase the amount of data. PointNet proposed by Charles et al. is the first deep learning network that directly takes point cloud as input without data conversion of point cloud. Figure 1 shows the network structure diagram of PointNet semantic segmentation model.

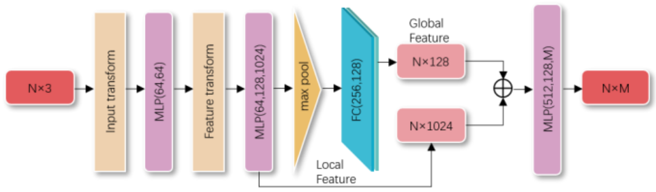

Figure 1: PintNet semantic segmentation network structure diagram

In the semantic segmentation network of PointNet, the input of the network is an $N \times 3$ point cloud matrix, where N is the total number of input sample points, and 3 represents the (X, Y, Z) three-dimensional space coordinates of each point. The network uses the shared weight MLP to increase the dimension multiple times, and extracts the features of each point to obtain the high-dimensional local features of the point, and then maps the extracted high-dimensional features to one dimension through the Max Pooling layer to obtain the point. Global characteristics of clouds. The segmentation network further enriches the features of points by aggregating local features and global features. Finally, through the MLP as a classifier, the features of the point cloud are mapped to the category interval, and the score of each category is obtained, where M represents the number of categories of semantic segmentation.

The loss function used by PointNet is the cross entropy loss function (The Cross Entropy Loss, CE Loss). Its mathematical expression is shown in formula (1):

$$Loss_{CE} = -\frac{1}{N}\sum_{i} L_i = -\frac{1}{N}\sum_{i}\sum_{c=1}^{M} y_{ic}\, log\left(p_{ic}\right) \qquad (1)$$

Among them, y_{ic} represents the sign function (0 or 1), and takes 1 when the real class of the sample point i is c, otherwise takes 0; p_{ic} represents the predicted probability that the observed sample point i belongs to the class c.

2.2 Sample equalization

The CE Loss used by PointNet has an obvious disadvantage. When using a dataset with unbalanced samples for training, although the loss of a large number of simple samples is small in single value, it forms the main part of the loss value through continuous superposition and dominates the gradient. The loss of difficult samples only accounts for a small part of the total loss value, which limits the recognition of difficult samples by the network. However, in brain semantic segmentation, difficult samples with smaller partitions do not mean their importance is low. For this reason, a sample balance mechanism needs to be introduced to coordinate the problem of uneven partitioning.

The mathematical model of the loss function Focal Loss of the sample equalization mechanism used in this paper is shown in formula (2):

$$Loss_{FL} = -\frac{1}{N}\sum_{i}\sum_{c=1}^{M}\alpha_c (1-p_{ic})^{\gamma} y_{ic} \log\left(p_{ic}\right) \qquad (2)$$

$$\alpha_c = 1/(\frac{n_c}{N}) \qquad (3)$$

In this method, the class modulation factor α_c and the modulation factor $(1-p_{ic})^{\gamma}$ are added on the basis of CE Loss to alleviate the problem of sample imbalance. Among them, the class modulation factor α_c is shown in formula (3), depending on the ratio of the c class sample points to the total sample points, so that the small partition samples have larger weights during training, and the large partition samples use smaller weights . The modulation factor $(1-p_{ic})^{\gamma}$ represents the weight of the dynamic correction for the samples of the ccc class partition after each training, which can make the low-accuracy samples in the current iteration obtain larger weights, where γ is a hyperparameter: a tunable focusing parameter. The two factors work together to determine the true weight of the c class samples in training, and jointly achieve the purpose of sample balance.

3. Experiment

3.1 Experimental data

The data set used in this experiment is MindBoggle101 data set [20] (M101 data set), which includes 5 sub-data sets, a total of 101 subjects' 3D brain MRI data, all data are reconstructed, and the cerebral cortex is obtained. point cloud collection.

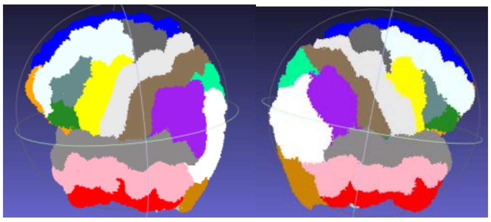

Figure 2: The example of MindBoggle101 dataset

Each sample in the dataset contains about 570, 000 to 750, 000 points. Due to the large size of the data, it cannot be input into the segmentation network at one time, and it needs to be processed in blocks.

3.2 Experimental environment and parameter settings

The experiments in this paper are all carried out on the Ubuntu 20.04 operating system, the graphics card uses NVIDIA/GeForce RTX 3080, the memory is 32G, the basic model uses the Tensorflow framework version of PointNet semantic segmentation network, the model training is performed with aaa learning rate, the Batch Size is set to 20, and the number of training times For 200 times, the experimental data were subjected to 5-fold cross-validation to test the robustness of the model.

3.3 Experimental evaluation indicators

In order to better verify the segmentation effect of the model and the ability of the model to deal with unbalanced samples, this paper adopts the Overall Accuracy (OA), mean Class Accuracy (mCA), the mean Intersection-Over-Union (mIoU) and the Standard deviation of Class Accuracy (SCA) were used as the evaluation indicators of the model.

OA represents the proportion of correctly classified points to the total number of points, mCA represents the average accuracy of the model on each semantic partition, OA and MCA reflect the overall segmentation ability of the model. The mathematical expressions of OA and mCA are shown in formula (4) and formula (5) respectively:

$$OA = \frac{n_{ii}}{N} \tag{4}$$

$$mCA = \frac{1}{c}\sum_{i=1}^{c}\frac{n_{ii}}{n_i} \tag{5}$$

The mIoU represents the average of the IoU of all categories, IoU represents the ratio of the intersection and union of the true value of the label and the predicted value, and mIoU considers the accuracy and recall of the segmentation results on all categories, which can measure the overall segmentation of the model. ability, its mathematical expression is as follows:

$$mIoU = \frac{1}{c}\sum_{i=1}^{c}\frac{n_{ii}}{n_i + \sum_{j=1}^{c} n_{ji} - n_{ii}} \tag{6}$$

SCA represents the standard deviation of the accuracy rate on each semantic partition, which reflects the ability of the model to alleviate the imbalance of the number of sample points in the partition. Its mathematical expression is as follows:

$$SPA = \sigma(MA) = \sigma(\frac{n_{ii}}{n_i}) \tag{7}$$

3.4 Experimental results and analysis

3.4.1 PointNet implements cerebral cortex semantic segmentation

The way of generating the input data of the PointNet network is to divide the coordinate block size in the X and Y directions as 1 and the step size as 0.5 into standard space blocks, and directly discard the blocks with less than 100 points in the block. The value (4096) is partially discarded at random.

The cerebral cortex is a curved surface structure composed of irregular sulci. If the data processing method of the PointNet network is directly applied, the distribution of the data in the divided blocks is seriously uneven. When using the M101 data set, due to the excessive number of points in some blocks, Far exceeding the rated value, it will cause the loss of valid data in a large number of blocks. If the block size is adjusted to solve this problem, the number of blocks will be too large, and there will be a lot of redundancy in the data, which exceeds the amount of data that the GPU can handle.

In view of the above problems, the M101 data set is divided into standard space blocks from the coordinates of X, Y, and Z directions, the block size is adjusted to 20, the step size is 10, and the rated value of the number of points in the block is set to 5000. At this time, the rated value is greater than 90 % of the number of points contained in the block, the maximum number of points in the block does not exceed 120% of the rated value.

After processing the input data format, the PointNet network is applied to perform semantic segmentation on the M101 dataset, and the segmentation results are shown in Table 2.

Table 1: Experimental results directly applied to the PointNet network

Method	OA	mCA	SCA	mIoU
PointNet	0.774	0.744	0.102	0.369

Compared with the whole brain segmentation method, the traditional cerebral cortex segmentation network uses the original MRI image as the input of the network, and the M101 dataset is also one of its commonly used datasets. When the MRI image is directly used as the input data of the network, the data volume of a single sample can reach 7.22 million, while when using the method in this paper, the data volume of a single sample is only 570, 000-750, 000, and the difference in data volume between the two is nearly 10 times.

The SLANT-27 network used by the whole-brain segmentation method [5] occupies more than 11GB of memory, and the training takes 648 hours (27 days); while the PointNet network used in this paper only occupies 4GB of memory during training, and the whole training process takes only 648 hours (27 days). About 20 hours. Therefore, the method in this paper has great advantages in time complexity and space complexity.

Table 2: Comparison of PointNet and SLANT-27 network efficiency

Network	The amount of data	Memory usage	Training time
SLANT-27	7.22 million	11GB	648h
PointNet	0.57-0.75 million	4GB	20h

Although the current segmentation results seem ideal, further statistics on the results of each partition of the cerebral cortex will reveal the problem. The partition with the least average number of points in the M101 dataset is only 7806 points, while the partition with the most points has 328242 points, a difference of more than 40 times.

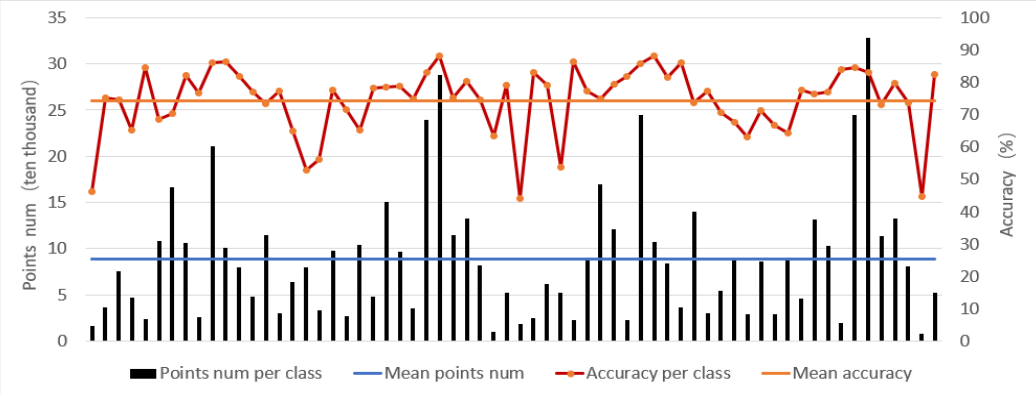

Figure 3: The accuracy of different partitions in the M101 dataset

Figure 3 shows the specific number of point clouds and segmentation accuracy for different partitions, where the orange straight line represents the average partition accuracy mCA. The segmentation results of the smallest and largest 10 semantic partitions in the statistical dataset are shown in Table 4. It can be seen that the results of semantic segmentation of larger partitions are obviously better than those of smaller partitions. After analysis, it is found that the main reason is that the samples of small partitions are insufficient, which will be overwhelmed by samples of large partitions during training and cannot be effectively trained, which ultimately affects Split effect.

Table 3: Segmentation results of min and max 10 partitions

Data set	mCA
Min 10 class	0.694
Max 10 class	0.807

3.4.2 Influence of sample balance mechanism on experimental results

Aiming at the problem that there is a large difference in the partition accuracy, the sample equalization mechanism is integrated into the PointNet network, and the same experiment is performed again. In the sample equalization mechanism, the hyperparameters of the Focal Loss loss function mainly include the class modulation factor α_c and the tunable focusing parameter γ. The value of α_c depends on the proportion of points in the current partition c to the total number of points in the sample. The value of γ is more flexible. During the experiment, the initial value is 1 and the step size is 1 for the experiment. When the value is 5, the experimental effect has no further improvement trend. Continue to select two adjacent γ parameter values with better experimental results in the existing experimental results to improve the γ value accuracy, and choose a step size of 0.1 for the experiment. The specific experimental results are shown in Table 4:

Table 4: Experimental results of M101 dataset under different loss function parameters

Loss Func	α	γ	OA	mCA	SCA	mIoU
CE Loss	×	×	0.774	0.744	0.102	0.369
Focal Loss	√	1	0.747	0.783	0.094	0.387
	√	1.1	0.747	0.783	0.092	0.389
	√	1.2	0.749	0.782	0.094	0.409
	√	1.3	**0.749**	**0.784**	**0.093**	**0.397**
	√	1.4	0.748	0.782	0.094	0.404
	√	1.5	0.746	0.782	0.093	0.411
	√	1.6	0.747	0.783	0.094	0.397
	√	1.7	0.747	0.783	0.094	0.415
	√	1.8	0.746	0.781	0.097	0.419
	√	1.9	0.745	0.783	0.095	0.428
	√	2	0.745	0.781	0.098	0.388
	√	3	0.745	0.781	0.100	0.403
	√	4	0.745	0.780	0.095	0.422
	√	5	0.743	0.778	0.097	0.421

Combining various evaluation indicators, among the hyperparameter combinations with different γ values, the best experimental results are obtained when the γ value is 1.3. The overall accuracy rate of the sample is 74.92%, and the average accuracy rate of the partition is 78.47%, an increase of 4.04%. %.

The experimental results show that after adding the Focal Loss loss function considering the sample balance mechanism, the mCA value of the two groups of data has been significantly improved, and the SCA value has been significantly reduced, indicating that the model can effectively alleviate the problem of unbalanced samples in the M101 data set. However, the mIoU value in the experimental results did not change significantly, and the OA value decreased slightly, indicating that the addition of the sample equalization mechanism still improves the overall segmentation effect of the model.

After statistics, among the 64 semantic partitions after using the sample equalization mechanism, the accuracy of 38 categories has been significantly improved, and the accuracy of 20 regions has increased by more than 10%.

Figure 4 shows the accuracy of the smallest 10 categories in the M101 dataset when using CE Loss and the Focal Loss loss function using the combination of the optimal hyperparameters. It can be seen that the accuracy of the small categories in the dataset has been significantly improved. promote.

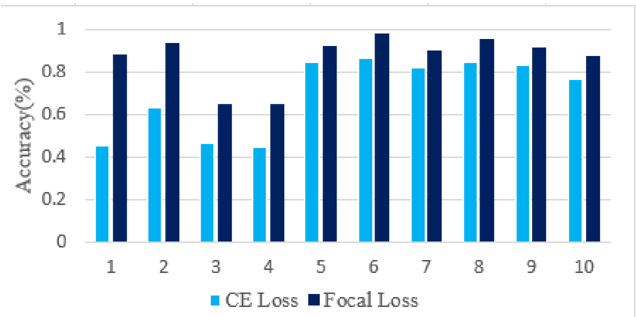

Figure 4: Accuracy of the 10 smallest partitions on the M101 dataset

4. Conclusion

In this paper, the PointNet semantic model is applied to the MindBoggle101 dataset to achieve the semantic segmentation of the cerebral cortex point cloud. At the same time, for the data imbalance problem of the cerebral cortex data in the semantic partition, the Focal Loss, which adjusts the sample balance mechanism, is introduced as the loss function of the model. So that the accuracy of the small sample partition of the model has been significantly improved. The experimental results are quantified by using indicators such as the overall accuracy of the sample, the average accuracy of the partition, the standard deviation of the accuracy of the partition, the average intersection ratio and the visual segmentation result of the sample. Moreover, the use of the sample balance mechanism can further improve the accuracy of semantic segmentation of small partitions. At the same time, this paper also compares with traditional segmentation methods, which proves that the method used in this paper can achieve more efficient segmentation of brain semantic partitions.

References

[1] Tarkhaneh O, Shen H. An adaptive differential evolution algorithm to optimal multi-level thresholding for MRI brain image segmentation [J]. Expert Systems with Applications, 2019, 138.

[2] Renugambal A, Bhuvaneswari KS. Image segmentation of brain MR images using Otsu's based hybrid WCMFO algorithm [J]. Computers Materials & Continua, 2020, 64 (2), 681-700.

[3] Lee B, Yamanakkanavar N, Choi J Y. Automatic segmentation of brain MRI using a novel patch-wise U-net deep architecture [J]. PLoS ONE, 2020, 15 (8).

[4] Ramzan F, Khan M U G, Iqbal S, Saba T, Rehman A. Volumetric segmentation of brain regions from MRI scans using 3D convolutional neural networks. IEEE Access 2020, 8, 103697-103709.

[5] Yuankai H, Zhoubing X, Yunxi X, Katherine A, Prasanna P, et al. 3d whole brain segmentation using spatially localized atlas network tiles [J]. NeuroImage. 2019, 194:105-119.

[6] Henschel L, Conjeti S, Estrada S, Diers K, Fischl B, Reuter M. FastSurfer - A fast and accurate deep learning based neuroimaging pipeline [J]. NeuroImage. https://doi.org/10.1016/j.neuroimage.2020.1170

[7] Pham T X, Siarry P, Oulhadj H. Integrating fuzzy entropy clustering with an improved PSO for MRI brain image segmentation. Applied Soft Computing, 2018, 65, 230-242.

[8] Boulanouar S, Lamiche C. A new hybrid image segmentation method based on fuzzy c-mean and modified bat algorithm. Int. J. Compute. Digit. Syst, 2020, 9, 677-687.

[9] Yang X, Xia D, Kin T and Igarashi T. 2020. IntrA: 3D intracranial aneurysm dataset for deep learning [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020:2653-2663

[10] Gutiérrez-Becker B and Wachinger C. Deep multi-structural shape analysis: application to neuroanatomy [C].The 21st International Conference on Medical Image Computing and Computer Assisted Intervention, 2018 523-531

[11] Ghazvinian Z F, Anssari M D, Verheij B, Claessen F, Cherici T and Tan T. 2019. Deep learning approach to semantic segmentation in 3D point cloud intra-oral scans of teeth [C]// International Conference on Medical Imaging with Deep Learning, 2019:557-571.

[12] Charles R. Qi, Hao Su, Kaichun Mo, et al. Pointnet: deep learning on point sets for 3d classification and segmentation [C] //the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 652-660.

[13] Ishikawa Y, Hachiuma R, Ienaga N, Kuno W, Sugiura Y, Saito H. Semantic Segmentation of 3D Point Cloud to Virtually Manipulate Real Living Space [C]. Asia Pacific Workshop on Mixed and Augmented Reality (APMAR), 2019: 63-69.

[14] Nurunnabi A, Teferle F N, Li J, Lindenbergh RC, Parvaz S. Investigation of PointNet for segmentation of largecscale outdoor point clouds [C]. 6th International Conference on Smart City Applications, 2021: 397-404.

[15] Mirsu R, Simion G, Caleanu C D, Pop-Calimanu I, M.A PointNet-Based Solution for 3D Hand Gesture Recognition [J]. SENSORS, 2020, 20 (7).
[16] Zhang L, Wang H. A novel segmentation method for cervical vertebrae based on PointNet plus plus and converge segmentation [J]. Computer Methods and Programs in Biomedicine, 2021, 200.
[17] Shrivastava A, Gupta A, Girshick R. Training region-based object detectors with online hard example mining [C] //the IEEE conference on Computer Vision and Pattern Recognition, 2016: 761-769.
[18] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017: 2980-2988.
[19] Li B, Liu Y, Wang X. Gradient harmonized single-stage detector [C] //the AAAI Conference on Artificial Intelligence. 2019, 33 (01): 8577-8584.
[20] Klein A, Mensh B, Ghosh S, et al. Mindboggle: automated brain labeling with multiple atlases [J]. BMC medical imaging, 2005, 5 (1): 1-14.

Author Profile

Yanting Sang is studying at the School of Control and Computer Engineering, North China Electric Power University, the main research directions are deep learning and medical image processing.