Report of CADA-AS competition*

Luo Chunhua[1,3], Huang Bo[1,3], Gao Haojun[2,3], Xie Yufeng[1,3], Liu Heng[1,3], and Jiang Siqing[2,3]

[1] College of Computer Science and Technology, Zhejiang University
[2] College of Public Health, Zhejiang University
[3] Real Doctor AI Research Centre, Zhejiang University

Abstract. The brain is an important organ of the body, which can control and regulate the function of various parts of the body. While cerebral aneurysms, as one of the most common aneurysms, can cause irreversible damage to the brain and even death. Once they are found, surgery is the widespread treatment. Therefore, accurate segmentation of cerebral aneurysms is of great significance to surgeons performing cerebral aneurysms. To this end, this paper proposes an effective method for segmentation of the cerebral aneurysms for the MICCA2020 competition. Specifically, our method uses 3D U-Net of nnU-Net to segment the cerebral aneurysms. After validation, the performance of our method reaches a good level under the given metric.

Keywords: 3D U-net · nnU-Net · aneurysm · segmentation.

1 CADA-AS

In the segmentation task, we choose nnU-Net as the basic model framework.nnU-Net is a segmentation framework, which can adapt to the previously invisible data set without any user interaction, and has good results.

1.1 Model architecture

Considering that the aneurysm is an outward bulging of blood vessel wall, the spatial position of the vessel is of great significance for the accurate localization and segmentation of the aneurysm. So we choose 3D U-Net of nnU-Net as the skeleton structure: it replace ReLU activation functions with leaky ReLUs; use instance normalization instead of the more popular batch normalization; use 30 feature maps at the highest resolution layers; start with a base configuration of input patch size 128 x 128 x 128, and a batch size of 2; pool (for a maximum of 5 times) along each axis until the feature maps have size 8.

1.2 Data preprocessing

Cropping All data is cropped to the region of nonzero values.

* Supported by Real Doctor AI Research Centre, Zhejiang University

Resampling All samples are resampled to the median voxel spacing of their respective dataset, where third order spline interpolation is used for image data and nearest neighbor interpolation for the corresponding segmentation mask.

Normalization All intensity values occurring within the segmentation masks of the training dataset are collected and the entire dataset is normalized by clipping to the [0.5, 99.5] percentiles of these intensity values, followed by a zscore normalization based on the mean and standard deviation of all collected intensity values.

1.3 Training

Training Schedule In our training process, we use the 5-fold cross validation method, which greatly increases the training time of the model, but helps to reduce the overfitting of the model. The network has converged after the training of 200 epochs. Each epoch contains 250 mini-batches which are randomly selected from training samples. In the meanwhile, oversampling is adopted to reduce the instability caused by random sampling. In term of model optimization, based on our previous experience, we adopt 0.01 as the initial learning rate, and the nesterov momentum rule is also used to make sure model converges faster. The decay of learning rate is carried out according to poly learning rate policy. In addition, data enhancement is crucial for the model to achieve good results. In the whole training process of the model, we use dynamic data enhancement, such as rotation, scaling, Gaussian noise, Gaussian blur, brightness, gamma, and so on , to increase the diversity of samples and make the model have better generalization ability.

Loss We adopt the loss function combined by the cross entropy loss and dice loss, to increase the stability of training and segmentation accuracy. and the results are improved to some extent. Because we're using a loss function like this, the best loss we can get is -1 and the loss will go down over the course of training.

$$L_{total} = L_{dc} + L_{CE} \tag{1}$$

$$L_{dc} = -\frac{2}{|K|} + \sum_{k \in K} \frac{\sum_{i \in I} u_i^k v_i^k}{\sum_{i \in I} u_i^k + \sum_{i \in I} v_i^k} \tag{2}$$

Experiments environment We conduct all our experiments on five TITAN Xp GPUs using the Pytorch 1.1 deep learning platform. The training process ends at 200 epochs, which takes approximately 27 hours.

Result Here is a demonstration of the results of the 5-fold cross validation process. The loss value curve and evaluation value curve of the five training processes are shown as follows which all have the same shape, and converge within 200 epochs. Every training has a good global average foreground dice.

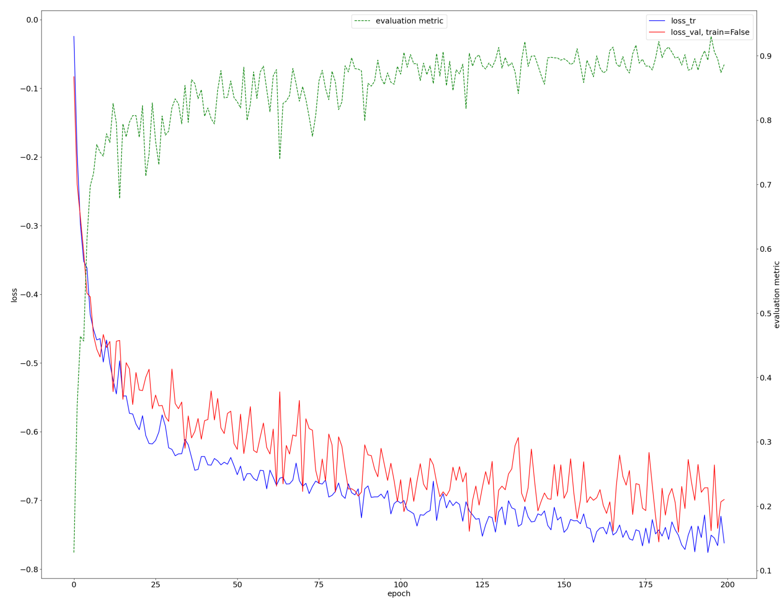

Fig. 1. the loss curve and evalution metric curve of one training process.

1.4 Inference

We trian U-Net with 3d full resolution configuration in a 5-fold cross-valiation, then use 5 models resulting from the cross-validation for ensembling.Cases are predicted using a sliding window approach.First, the step size of sliding window is 0.5, which means half the patch size.Second, data augmentation in the form of mirroring is done along all three axes.Third, we uses a Gassuian importance weighting to weigh predictions closer to the center of the current patch higher than those at the borders as the segmentation accuracy decreases towards the borders.

References

1. F., Petersen, J., Klein, A., Zimmerer, D., Jaeger, P.F., Kohl, S., and, ...: nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation (2018). arXiv:1809.10486.
2. Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., ... Glocker, B.: Attention u-net: Learning where to look for the pancreas (2018). arXiv:1804.03999.
3. Ronneberger, O., Fischer, P., Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.