Intracranial Hemorrhage Segmentation method based on nnU-Net

Jinhuang Zhang[1] and Xiangyuan Ma[1]

[1] Shantou University, Department of Biomedical Engineering, China
21jhzhang1@stu.edu.cn

Abstract. Intracranial hemorrhage is a common severe brain disease in the process of neurosurgery, which has the characteristics of severe disease, rapid change, high mortality and disability rate. CT is the fastest and most convenient tool to diagnose ICH, and is also the most commonly used technique in emergency department. In addition to providing a definite diagnosis, CT can also show the basic features of the hematoma. Craniocerebral CT imaging data of intracranial hemorrhage is huge, and the diagnosis accuracy is high. If it can make a breakthrough in computer-aided diagnosis, it will help doctors to make clinical diagnosis and treatment and adjust the optimal treatment plan. In this paper, we used the typical nnU-Net to do the task of Intracranial Hemorrhage Segmentation Challenge on Noncontrast head CT (Named INSTANCE 2022). Finally, based on the trained network, the mean dice of the prediction results come to 0.73386.

Keywords: Intracranial Hemorrhage, Segmentation, nnU-Net.

1 Introduction

Intracranial hemorrhage is a common severe brain disease in the process of neurosurgery, which has the characteristics of severe disease, rapid change, high mortality and disability rate[1,2]. Bleeding in the brain CT images there is no mature scheme of computer intelligent recognition, we use a variety of computer models as well as the method of image segmentation, intends to explore the effective bleeding in the brain CT image recognition, artificial intelligence in the future early diagnosis of cerebral hemorrhage, auxiliary application treatment, improve the prognosis of patients, have very positive meaning.

At present, the most important imaging examination in the diagnosis of intracranial hemorrhage is brain CT plain scan, which has the characteristics of fast, clear and accurate imaging. The study of imaging results in patients with intracranial hemorrhage is helpful to assist the diagnosis and adjust the best treatment plan. Craniocerebral CT imaging data of intracranial hemorrhage is huge, and the diagnosis accuracy is high. If it can make a breakthrough in computer-aided diagnosis, it will help doctors to make clinical diagnosis and treatment and adjust the optimal treatment plan.

Therefore, it is necessary to establish a fully-automated segmentation method, which allows accurate and rapid volume quantification of the intracranial hemorrhage. However, it is still challenging to accurately segment the ICH for automatic methods because ICH exhibits large variations in shapes and locations, and has blurred boundaries. [3]
Different methods are used to do the task, which can be generally encouraging.

2 Methods

nnU-Net[4] pays more attention to parts outside the structure, such as preprocessing, training and reasoning strategies, and post-processing.

2.1 Network architectures

Medical images commonly encompass a third dimension, which is why a pool of basic U-Net architectures consisting of a 2D U-Net, a 3D U-Net and a U-Net Cascade are considered. While the 2D and 3D U-Nets generate segmentations at full resolution, the cascade first generates low resolution segmentations and subsequently refines them. The architectural modifications as compared to the U-Net's original formulation are close to negligible and instead focus our efforts on designing an automatic training pipeline for these models.

The U-Net[5] is a successful encoder-decoder network that has received a lot of attention in the recent years. Its encoder part works similarly to a traditional classification CNN in that it successively aggregates semantic information at the expense of reduced spatial information. Since in segmentation, both semantic as well as spatial information are crucial for the success of a network, the missing spatial information must somehow be recovered. The U-Net does this through the decoder, which receives semantic information from the bottom of the 'U' and recombines it with higher resolution feature maps obtained directly from the encoder through skip connections. Unlike other segmentation networks, such as FCN[6] and previous iterations of DeepLab[7] this allows the U-Net to segment fine structures particularly well.

Just like the original U-Net, nnU-Net use two plain convolutional layers between poolings in the encoder and transposed convolution operations in the decoder. We deviate from the original architecture in that this replaces ReLU activation functions with leaky ReLUs (neg. slope 1e−2) and uses instance normalization instead of the more popular batch normalization.

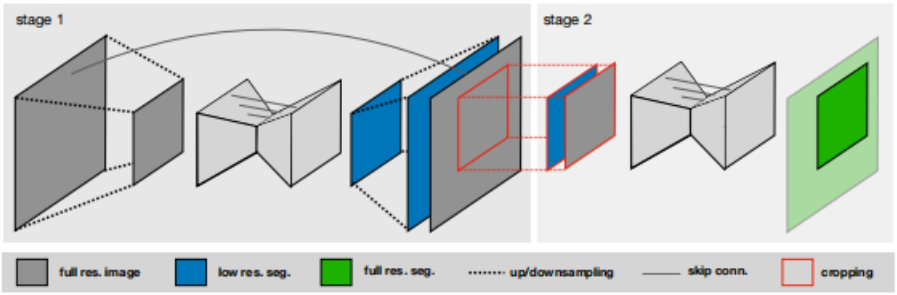

Fig. 1. U-Net Cascade (on applicable datasets only). Stage 1 (left): a 3D U-Net processes downsampled data, the resulting segmentation maps are upsampled to the original resolution. Stage 2 (right): these segmentations are concatenated as one-hot encodings to the full resolution data and refined by a second 3D U-Net.

2.2 Preprocessing

The preprocessing is part of the fully automated segmentation pipeline that our method consists of and, as such, the steps presented below are carried out without any user intervention.

Cropping All data is cropped to the region of nonzero values. This has no effect on most datasets such as liver CT, but will reduce the size (and therefore the computational burden) of skull stripped brain MRI.

Resampling CNNs do not natively understand voxel spacings. In medical images, it is common for different scanners or different acquisition protocols to result in datasets with heterogeneous voxel spacings. To enable our networks to properly learn spatial semantics, all patients 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.

Necessity for the U-Net Cascade is determined by the following heuristics: If the median shape of the resampled data has more than 4 times the voxels that can be processed as input patch by the 3D U-Net (with a batch size of 2), it qualifies for the U-Net Cascade and this dataset is additionally resampled to a lower resolution. This is done by increasing the voxel spacing (decrease resolution) by a factor of 2 until the mentioned criterion is met. If the dataset is anisotropic, the higher resolution axes are first downsampled until they match the low resolution axis/axes and only then all axes are downsampled simultaneously. The following datasets of phase 1 fall within the set of described heuristics and hence trigger usage of the U-Net Cascade: Heart, Liver, Lung, and Pancreas.

Normalization Because the intensity scale of CT scans is absolute, all CT images are automatically normalized based on statistics of the entire respective dataset: If the modality description in a dataset's corresponding json descriptor file indicates 'ct', 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 z-score normalization based on the mean and standard deviation of all collected intensity values. For MRI or other image modalities (i.e. if no 'ct' string is found in the modality), simple z-score normalization is applied to the patient individually.

2.3 Training

For 3D U-Nets operating on nearly entire patients (first stage of the U-Net Cascade and 3D U-Net if no cascade is necessary) we compute the dice loss for each sample in the batch and average over the batch. For all other networks we interpret the samples in the batch as a pseudo-volume and compute the dice loss over all voxels in the batch.

2.4 Inference

Due to the patch-based nature of our training, all inference is done patch-based as well. Since network accuracy decreases towards the border of patches, we weigh voxels close to the center higher than those close to the border, when aggregating predictions across patches. Patches are chosen to overlap by patch size / 2 and we further make use of test time data augmentation by mirroring all patches along all valid axes.

Combining the tiled prediction and test time data augmentation result in segmentations where the decision for each voxel is obtained by aggregating up to 64 predictions (in the center of a patient using 3D U-Net). For the test cases, the five networks obtained from training set cross-validation are used as an ensemble to further increase the robustness of the models.

3 Experiment and result

3.1 Dataset

The dataset[8,9] used is collected by the professors in Intracranial Hemorrhage Segmentation Challenge on Noncontrast head CT (Named INSTANCE 2022), which will be served as a solid benchmark for Intracranial hemorrhage Segmentation tasks.

In details, 200 3D volumes with refined labeling from 10 experienced radiologists have been collected, 100 for the training dataset, 70 for the closed testing dataset, and 30 for the opened validated dataset. DSC, HD, RVD are adopted as evaluation metrics for segmentation. This challenge will also promote intracranial hemorrhage treatment, interactions between researchers, and interdisciplinary communication.

3.2 Experiment process

First, we put the data into the corresponding folder created in the format, The nnUNet_convert_decathlon_task command is used to automatically convert the format.
Then, we choo se 3D full resolution U-net to fit the task. The 3D full resolution U-net has been trained for 1000 epoch. Blow are some parameters settings:
Training process -- training from scratch, using 5 fold cross-validation, loss function: combining DICE loss and cross-entropy loss:

For 3D-UNets trained on the full training set (the first stage of UNet Cascade and the non-cascaded 3D UNet, excluding the second stage of UNet Cascade), DICE loss is calculated separately for each sample and then averaged at batch. For other networks (2D UNet and the second phase of UNet Cascade), DICE on the whole batch is calculated by treating all samples in a batch as a whole sample.

Adam optimizer, learning rate 3E-4; 250 batch/epoch;

Learning rate adjustment strategy: Calculate the exponential moving average loss of the training set and the validation set. If the exponential moving average loss of the training set is not reduced by 5E-3 within 30 epochs, the learning rate will decrease by 5 times.

Training stop conditions: When the exponential moving average loss of the validation set is less than 5E-3 within 60 epochs, or the learning rate is less than 1E-6, the training will be stopped

Data augmentation: random rotation, random scaling, random elastic transformation, gamma correction, mirror.

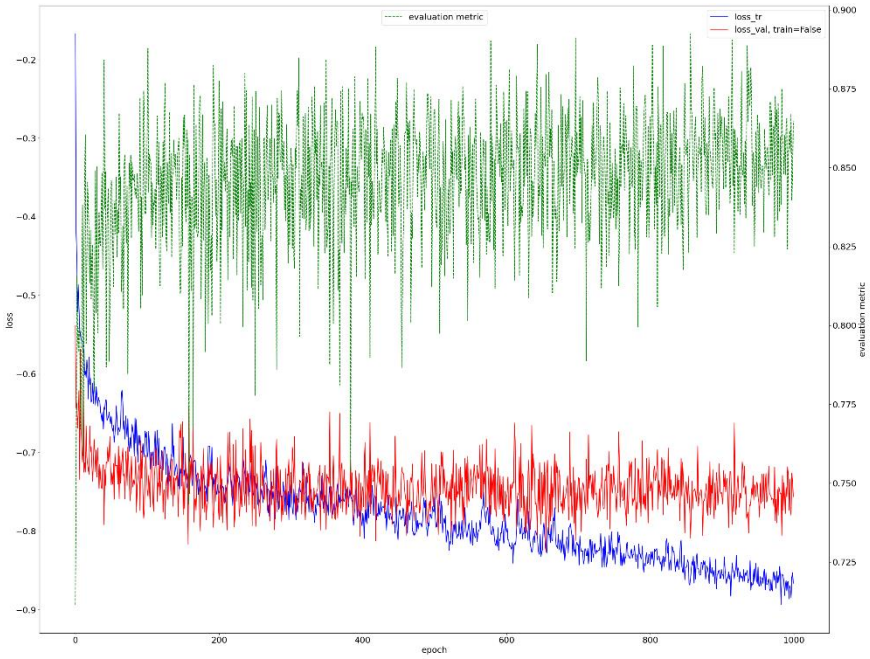

Fig. 2. Loss and dice figure during the training process. Blue for training loss, red for validation loss, and green for dice.

3.3 Result

Finally, 30 3D volumes are used for the opening validation phase. Based on the trained network, we get the prediction results, whose mean dice values 0.73386, and ranks 10th place in the validation place.


```
"DiceCoefficient": {
  "max": 0.9452558923199176,
  "min": 1.1086474378198733e-08,
  "std": 0.27383900290640034,
  "25pc": 0.7051770403409132,
  "50pc": 0.870306999339498,
  "75pc": 0.9006122723416847,
  "mean": 0.7338595817137015,
  "count": 30.0
},
```


Fig. 3. Dice values of predictions in validation phase

References

1. C. J. van Asch, M. J. Luitse, G. J. Rinkel, I. van der Tweel, A. Algra, and C. J. Klijn, “Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: A systematic review and meta-analysis,” Lancet. Neurol., vol. 9, no. 2, pp. 167–176, Feb. 2010.
2. J. N. Goldstein and A. J. Gilson, “Critical care management of acute intracerebral hemorrhage,” Curr. Treat. Option. Neurol., vol. 13, no. 2, pp. 204–216, Jan. 2011.Author, F., Author, S., Author, T.: Book title. 2nd edn. Publisher, Location (1999).
3. INSTANCE grand-challenge https://instance.grand-challenge.org/
4. F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, “No new-net,” in International MICCAI Brainlesion Workshop. Springer, 2018, pp.234–244.
5. O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in MICCAI. Springer, 2015, pp. 234–241.
6. J. L. E. S. T. Darrell and U. Berkeley, “Fully Convolutional Networks for Semantic Segmentation,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR). Boston, MA, USA IEEE, 2015. 3431~3440., pp. 847–856, 2015, doi: 10.1109/ICCVW.2019.00113.
7. L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution,and fully connected crfs,” IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 4, pp. 834–848, 2018.
8. X. Li, G. Luo, W. Wang, K. Wang, Y. Gao and S. Li, "Hematoma Expansion Context Guided Intracranial Hemorrhage Segmentation and Uncertainty Estimation," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 3, pp. 1140-1151, March 2022, doi: 10.1109/JBHI.2021.3103850.
9. Xiangyu Li, Kuanquan Wang, Jinbo Liu, Hongyu Wang, Mingwang Xu, & Xinjie Liang. (2022). The 2022 Intracranial Hemorrhage Segmentation Challenge on Non-Contrast head CT (NCCT). 25th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2022).