3D Cascade U-Net for Intracranial Hemorrhage Segmentation on CT Images

Xingyu Zhao[1] [0000-0001-6594-8853] , Shan Huang[1], Minglei Liu[1],

Weiping Liu[1](✉), Jiayin Zhou[1]

[1] Shanghai Microport Medbot (Group) Co.,Ltd., China
weiping.liu@microport.com

Abstract. Intracranial hemorrhage (ICH) is one of the common stroke types. Diagnosis and risk evaluation of ICH is significant for patient, for which ICH segmentation is the first and key step to do. However, it is incredibly time-consuming and subjective, which may result in poor consistency and low efficiency of diagnoses. Therefore, a fully automatic segmentation method is urgently developed for ICH segmentation. In this paper, a 3D cascade U-Net network was adapted to carry out the ICH segmentation based on the CT images. We evaluated our model on the opened validated dataset by 5-fold cross validation experiments. With a final Dice Similarity Coefficient (DSC) of 0.734, a Hausdorff Distance (HD) of 27.03, a Relative Volume Difference (RVD) of 0.212, and a Normalized Surface Dice (NSD) of 0.510. This deep learning-based segmentation model can successfully achieve ICH segmentation on CT, and holds great potential in clinical practice.

Keywords: Intracranial hemorrhage, Automatic, Cascade.

1 Introduction

Intracranial hemorrhage (ICH) is a common stroke type and has the highest mortality rate among all stroke types [1], and Non-Contrast Computed Tomography (NCCT) is the most widely used modality for diagnosing ICH [2]. Diagnosis and risk evaluation of ICH is significant, for which intracranial hemorrhage segmentation is the first step [3]. Nowadays, the ICH regions were manually delineated by radiologists, which is time-consuming and suffers from inter-rater variability. Also, poor repeatability and susceptibility to human errors in segmentation might exist. Therefore, a more objective, precise, and convenient method should be developed for ICH segmentation.

In recent years, deep learning has been widely used in image classification, segmentation, and object detection [4,5] due to its high robustness and efficiency. In this paper, a cascaded 3D U-Net was adopted to segment ICH based on the CT images.

2 Data and Method

2.1 Dataset

The dataset we used in this paper are from the INtracranial hemorrhage SegmenTAtioN ChallengE (INSTANCE)[6,7]. Data from 100 patients with ICH used as the training dataset, and another 30 for the opened validated dataset. Each case includes CT in NIfTI format, as well as corresponding hemorrhage annotations by ten experienced radiologists.

2.2 Data Preprocessing

To improve the quality of the data so that the network can easily learn the features of the data, the following preprocessing operations are performed to the dataset of INSTANCE 2022 in this paper. First, owing to the different tissue and regions can be showed in different windows and levels, along with the morphological structure of the hemorrhage is strongly related to the position of the hematoma, the HU of CT images were clipped according to three different windows and levels, and corresponding range of HU were [0, 80], [-20, 180] and [-150, 230]. The intensity of the voxel above the range were assigned the value of upper limit in range, and the intensity below the range is assigned the value of lower limit in range. The image is shown in Fig. 1. Second, the three images with different HU range clip were served as three channels and treated as one image, which can better integrate the prior information of window and level of the ICH image for model training. The pseudo color image was shown in **Fig. 1**.

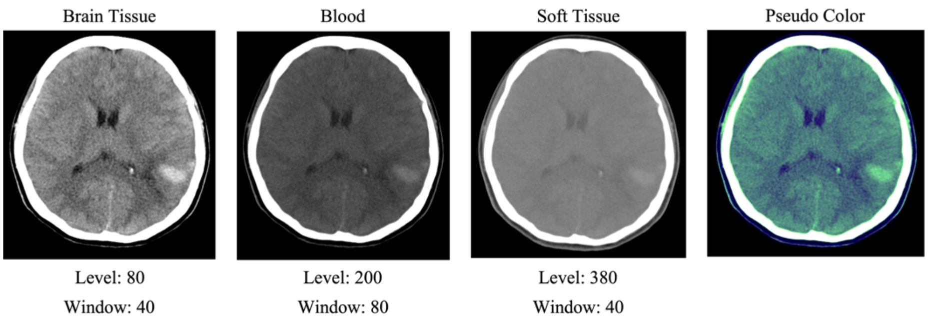

Fig. 1. Visual examples of CT with different HU windows and levels in first three images. The last one is a pseudo image generated by first three.

2.3 Network Structure

The network we adopted is two stage 3D U-Net [8] based cascade model. The architecture is shown in **Fig. 2.**

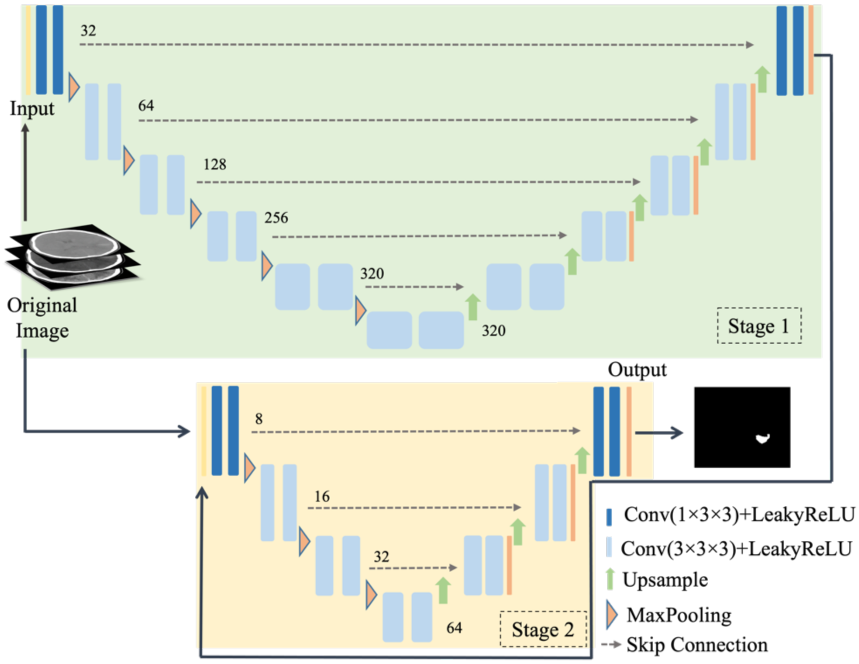

Fig. 2. The architecture of segmentation network.

For the Stage 1, the basic module of the encoder and decoder is Conv-Instance Norm-LeakyReLU [9]. The operation of downsampling in the encoder is achieved by max pooling. The upsampling operation in the decoder is achieved by using the transpose convolution of $2 \times 2 \times 2$. There are initial 32 feature maps in this network architecture, and they are doubled along each downsampling operation in the encoder, but the maximum feature maps is no more than 320. They will be halved by each transposed convolution in the decoder. At the end of the decoder, spatial size achieved is the same as initial input size. Then there is a $1 \times 1 \times 1$ convolution into 1 channel and a sigmoid function followed.

For the Stage 2, to further achieve fine segmentation and improve the segmentation performance of the model, a 3D U-Net was cascaded to the model, whose input is the output of probability map of the first stage.

2.4 Experiments

We used the cascade model defined in Section 2.3 as the baseline model. In addition, we evaluated the original image as the input to train the baseline. And the model without stage 2 was also trained and evaluated with different loss combination mode. The program is implemented using nnUNet [10].

The loss function used in experiments is the sum of Dice loss and Binary Cross Entropy loss, defined as:

$$L_{seg} = -\frac{2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} - \sum_i^N (p_i \log g_i + (1 - p_i)\log(1 - g_i)) \tag{1}$$

The training process is optimized using Adam optimizer [11] with an initial learning rate 0.01. The 5-fold cross-validation experiments are carried out. The given training data by the INSTANCE 2022 Challenge organization are randomly partitioned into 5 folds, and 4 folds from the 5 ones are used as the training data to train our model, and the remaining one as the evaluation fold. It took about 43 hours to train one model. There are 5 models achieved in total. The 5 models were evaluated on the opened validation dataset, and the mean results are recorded.

The experiments are implemented in the environment of Linux operating system, and the NVIDIA A100 PCIE with 40G RAM is used to accelerate training.

3 Result

The performance of the models was evaluated on the opened validated dataset, and corresponding performance of the models in Top three and their ensemble were shown in Table 1. Baseline outperformed other non-ensemble models using the same loss function for training. And the performance can further improve by integrated top three models in terms of DSC.

Table 1. Model evaluation on the opened validated dataset.

	DSC	HD	RVD	NSD
Baseline	0.7343	27.034	0.2119	0.5095
Baseline (1 Channel Input)	0.7236	30.1010	0.2708	0.5151
Baseline + BCE	0.7167	41.5849	0.2838	0.4775
Ensemble	0.7355	27.1629	0.2364	0.5157

4 Discussion

Table 1 shows that best results of non-ensemble model on the opened validated dataset data are 0.7343, 27.034, 0.2119 and 0.5095 in terms of Dice, HD, RVD and NSD, respectively. All of them were from Baseline model, which proves the 3-channel image served as input can be easier to train the model than the 1-channel original image. In addition, the ensemble model was generated by top 3 models, whose performance can be further improved in terms of DSC, which indicated that derived knowledge from different models are complementary.

Some segmentation results were shown in Fig. 3. For the first row, the image shows that the prediction of the model is very close to the ground truth, which proved that the model can achieve a good performance for some cases. However, for another case (the second row in Fig. 3), it cannot be well segmented. Based on the results, we infer that low contrast with surrounding tissues, where less gray features of images could be extracted to segment ICH.

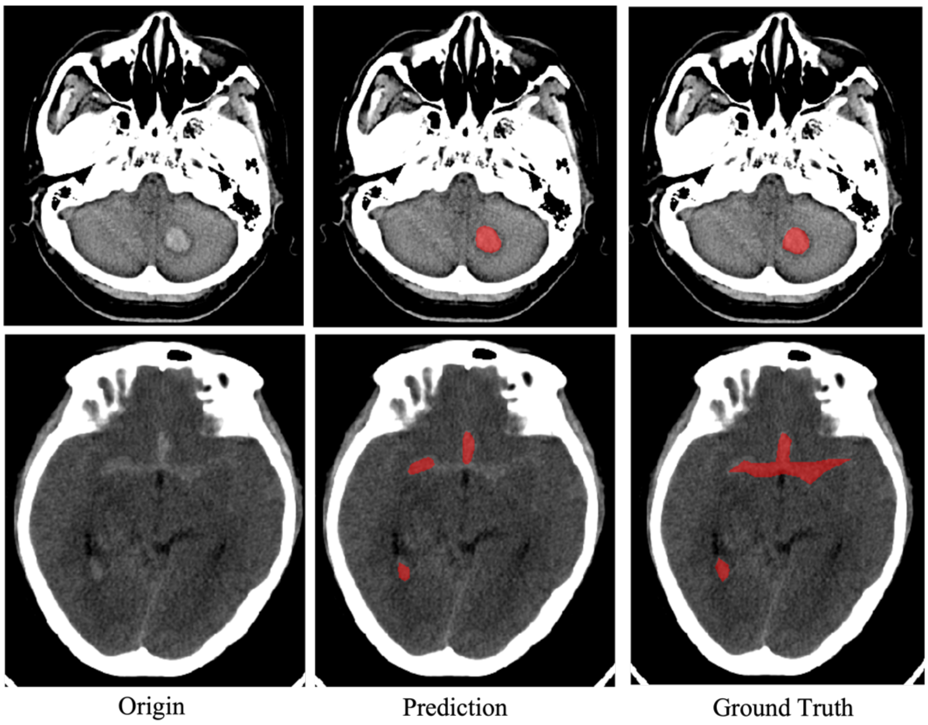

Fig. 3. The images in first column are original CT image; The second column are the model prediction results (red regions) overlayed on the original image; The third column shows the ground truth (red regions) overlayed on the original image.

5 Conclusion

In this paper we developed a fully automatic method using deeply supervised 3D U-Net based cascade network, which can take advantage of extracting more significant information, thus to effectively segmentate ICH. Although it works well, it still suffers from the insufficient generalization in some cases. We will continue to study how can solve this challenging work. And we also hope to work together with those scholars in the related research fields from all over the world to overcome the challenges in this field.

References

[1] Van Asch CJ, Luitse MJ, Rinkel GJ, van der Tweel I, Algra A, Klijn CJ. Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. The Lancet Neurology 2010;9:167–76.

[2] Goldstein JN, Gilson AJ. Critical care management of acute intracerebral hemorrhage. Current Treatment Options in Neurology 2011;13:204–16.

[3] Broderick JP, Brott TG, Duldner JE, Tomsick T, Huster G. Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality. Stroke 1993;24:987–93.

[4] Zhao X, Wang X, Xia W, Zhang R, Jian J, Zhang J, et al. 3D multi-scale, multi-task, and multi-label deep learning for prediction of lymph node metastasis in T1 lung adenocarcinoma patients' CT images. Computerized Medical Imaging and Graphics 2021;93:101987. https://doi.org/10.1016/j.compmedimag.2021.101987.

[5] Zhao X, Xie P, Wang M, Li W, Pickhardt PJ, Xia W, et al. Deep learning–based fully automated detection and segmentation of lymph nodes on multiparametric-mri for rectal cancer: A multicentre study. EBioMedicine 2020;56:102780. https://doi.org/10.1016/j.ebiom.2020.102780.

[6] Li X, Luo G, Wang W, Wang K, Gao Y, Li S. Hematoma Expansion Context Guided Intracranial Hemorrhage Segmentation and Uncertainty Estimation. IEEE Journal of Biomedical and Health Informatics 2022;26:1140–51. https://doi.org/10.1109/JBHI.2021.3103850.

[7] Li X, Wang K, Liu J, Wang H, Xu M, Liang X. The 2022 Intracranial Hemorrhage Segmentation Challenge on Non-Contrast head CT (NCCT) 2022. https://doi.org/10.5281/zenodo.6362221.

[8] Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation 2016. https://doi.org/10.48550/ARXIV.1606.06650.

[9] Zhang X, Zou Y, Shi W. Dilated convolution neural network with LeakyReLU for environmental sound classification. 2017 22nd international conference on digital signal processing (DSP), IEEE; 2017, p. 1–5.

[10] Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 2021;18:203–11. https://doi.org/10.1038/s41592-020-01008-z.

[11] Kingma DP, Ba J. Adam: A method for stochastic optimization. ArXiv Preprint ArXiv:14126980 2014.