nnU-Net for Intracranial Hemorrhage Segmentation

Chuanpu Li[1,2,3] and Zeli Chen[1,2,3]

[1] School of Biomedical Engineering, Southern Medical University, Guangzhou, China
[2] Guangdong Provincial Key Laboratory of Medical Image Processing, Southern Medical University, Guangzhou, China
[3] Guangdong Province Engineering Laboratory for Medical Imaging and Diagnostic Technology, Southern Medical University, Guangzhou, China

Abstract. Accurate segmentation of intracranial hemorrhage can save physicians time and provide volume estimation of the intracranial hemorrhage, which is essential for the diagnosis and treatment for patients. This paper describes our contribution to the Intracranial Hemorrhage Segmentation Challenge on Noncontrast head CT (Instance 2022). We developed our methods based on nnU-Net with little modification. Moreover, we developed an uncertainty estimation ensemble strategy. Experimental results with unseen validation data from the leaderboard showed the effects of our methodology.

Keywords: Intracranial Hemorrhage Segmentation · nnU-Net · Deep Learning.

1 Introduction

Intracerebral haemorrhage(ICH) is the second most common cause of stroke, and it is associated with high rates of fatality and functional disability[18,20]. In the clinical setting, accurate volume measurements of ICH plays a significant role in the prognosis and treatment decisions for patients[16]. Noncontrast computed tomography (CT) is the most widely used modality for hematoma assessment due to its rapid acquisition and widespread availability[3,5]. However, manually segmenting the ICH region in CT images by radiologists for hematoma volume estimation is both time-consuming[15] and suffers from inter-rater variability[8]. The ABC/2 formula, although widely used to estimate hematoma volume in clinical pratice due to its simplicity, often suffers from large estimation error especially for those irregular or lobar hematomas[22]. Therefore, the automated segmentation of intracranial hemorrhage can save physicians time and provide an accurate volume estimation of the intracranial hemorrhage.

Intracranial Hemorrhage Segmentation Challenge on Noncontrast head CT (INSTANCE) is aiming at being a solid benchmark for intracranial hemorrhage segmentation tasks and promoting intracranial hemorrhage treatment, interactions between researchers, and interdisciplinary communication[10]. This year, INSTANCE 2022 training dataset collected 100 non-constract head CT volumes of clinically diagnosed patients with different kinds of ICH, including subdural hemorrhage, epidural hemorrhage, intraventricular hemorrhage, intraparenchymal hemorrhage, and subarachnoid hemorrhage(see Fig. 1). The size of a CT slice is 512×512, and the number of slices lies in [20, 70]. The pixel spacing of a CT volume is 0.42mm $\times$ 0.42 mm $\times$ 5 mm. All of the CT volumes were manually segmented and cross-validated by 10 experienced neuro-radiologists. Two additional datasets without the labels were provided for validation and testing. The validation dataset (30 cases) allowed multiple submissions and was designed for intermediate evaluations. The testing dataset (70 cases) allowed only a single submission, and was used to calculate the final challenge ranking. A number of metrics(Dice score, Hausdorff distance and Relative Volume Difference) are used to measure the segmentation performance of the algorithms proposed by participants.

In this work, due to the effectiveness of nnU-Net in many datasets[7], we implemented the nnU-Net as the baseline and extended it to intracranial hemorrhage segmentation. Moreover, we proposed a simple but efficient uncertainty estimation ensemble strategy. Experimental results with unseen validation data from the leaderboard showed the effects of nnU-Net and our uncertainty estimation ensemble strategy.

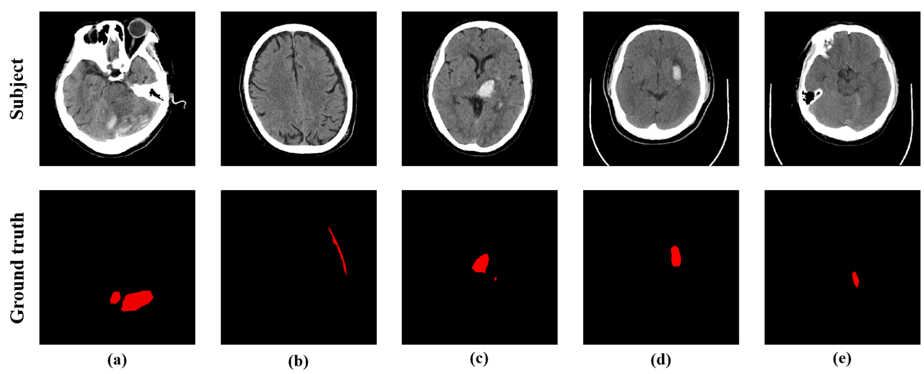

Fig. 1. Example of different kinds of intracranial hemorrhage in the INSTANCE 2022 dataset. (a)subdural hemorrhage (b)epidural hemorrhage (c)intraventricular hemorrhage (d)intraparenchymal hemorrhage and (e)subarachnoid hemorrhage.

2 Related Work

Stroke is currently the second leading cause of death in the world [14]. It is important to detect intracranial hemorrhage and segment using a computer-aided detection or diagnosis (CAD)-based automatic system.In the current methods, intracranial hemorrhage segmentation is mainly divided into traditional methods and deep learning methods.

Traditional methods Bhadauria *et al.* presented a segmentation technique that combines the features of fuzzy c-mean (FCM) clustering and region-based active contour method, which has advantages in accuracy in comparison with standard region growing method [2]. Muschelli *et al.* adopted a random forest algorithm with features extracted from X-ray computed tomography (CT) scans for automatic segmentation [13].In the actual modeling process, these traditional methods [1,2,4,13] often need tedious preparation work, such as image registration and skull stripping, and the segmentation performance and running time are not ideal.

Deep learning methods Deep learning technology has achieved good performance in many medical image tasks because of its powerful learning function [21,11]. At present, UNet is the most commonly used network structure in medical image segmentation [17]. It obtains the details and texture information of the image through the fusion of different feature maps and skip connection, which makes the network can learn more deeply and stronger. Ironside *et al.* [6] and Kwon *et al.* [9] utilized UNet in intracranial hemorrhage segmentation achieved a good performance. SLEX-Net [10] proposed incorporated hematoma expansion in the segmentation architecture by directly modeling the hematoma variation among adjacent slices to improve the accuracy and continuity of segmentation results. nnU-Net [7] is a self-adapting framework with powerful preprocessing, model framework and post-processing. It is an important reference method for current medical image segmentation, but it has a high integration degree and cannot be adjusted for specific tasks.

3 Methods

In this section, We first point out the challenge of intracranial hemorrhage segmentation. Then models and strategies for addressing these challenges are elaborated. All of our experiments were done with the excellent open-source nnU-Net framework.[4]

3.1 Challenges for Intracranial Hemorrhage Segmentation

Although there have been numerous deep-learning based methods(!!! add ref !!!) for intracranial hemorrhage segmentation, we find that it is still challenging for the following reasons: 1) As shown in Fig. 1, there are many types of intracranial hemorrhage and therefore many variations in shape and location. 2) Intracranial hemorrhage has blurred boundaries, and some even hard to distinguish from the bone(see Fig. 1(b)). 3) Unbalanced intracranial hemorrhage types and some small intracranial hemorrhage areas tend to confuse models, which makes the model to generate completely different results, even though they have the exact same architecture and just trained in different folds.

[4] https://github.com/MIC-DKFZ/nnUNet

3.2 Models and Strategies for Addressing the Challenges

Baseline nnU-Net Since the pixel spacing of the INSTANCE 2022 CT volume is 0.42mm × 0.42 mm × 5 mm, which has more detailed information in xy plane, we first tried the 2D nnU-Net. This 2D nnU-Net has the encoder-decoder structure with skip-connection operating on image size 512 × 512. The encoder contains 7 level of 2D convolutional layers with strided convolutional layers downsampling. The decoder follows the same structure with transpose convolution upsampling and convolution operating on concatenated skip features from the encoder branch at the same level. Leaky ReLU with slope of 0.01[12] and instance normalization[19] was applied after every convolution operations. Then we tried the 3D nnU-Net.The 3D nnU-Net has the similar structure with 2D nnU-Net but operates on patch of size 16 × 320 × 320 and only contains 5 level of 3D convolutional layers. Our experiments showed that even though the 2D nnU-Net could not achieve the overall dice accuracy of 3D nnU-Net, it performed better results than 3D nnU-Net when the intracranial hemorrhage had very small area or blurred boundaries. Therefore, it is necessary to use both 2D and 3D nnU-Net to predict the final result.

Loss Function In order to further alleviate the segmentation issue of small area intracranial hemorrhage and maintain stability during training, we utilized the weighted cross-entropy loss to replace simple cross-entropy loss in the nnU-Net, which is defined as:

$$L_{wce} = -\frac{1}{N}\sum_{n=0}^{N}\sum_{c=0}^{1} w^c y_n^c \log \hat{y}_n^c \tag{1}$$

where $\hat{y}_n^c$ denotes the probability from the model that pixel n belongs to class c ($c = 1$ for hemorrhage and $c = 0$ background), y_n^c is the label at pixel n, ω^c denotes the weight for class c. We set $\omega^0 = 0.2$ and $\omega^1 = 0.8$ in our experiments. And the final loss fuction is:

$$L = \lambda_1 L_{dice} + \lambda_2 L_{wce} \tag{2}$$

where L_{dice} indicates the dice loss. λ_1 is set to 0.6 and L_2 is set to 0.4.

3.3 Training Details

We followed the training procedure of nnU-Net for all networks. Each network was trained with 5-fold cross validation. A large variety of data augmentation techniques are applied, including random rotations, random scaling, random elastic deformations, gamma correction augmentation and mirroring. The networks were optimized with Adam with optimizer with an initial learning rate 0.01,and was decayed following a polynomial schedule:

$$lr = 0.01 \times \left(1 - \frac{epoch}{1000}\right)^{0.9} \tag{3}$$

Each training run lasted 1000 epochs, with each epoch consisting of 250 mini-batches. The dice score on the validation set of the current fold was used to monitor the training progress. All experiments were conducted with Pytorch 1.11 on NVIDIA RTX 2080 Ti GPU with 12GB VRAM.

3.4 Uncertainty Estimation Ensemble Strategy

Due to the unbalanced intracranial hemorrhage types and intracranial hemorrhage areas, the models trained in different folds might predict completely different results, especially for those cases which has blurred boundaries and small areas. Simply average the predicted results from the models provide no additional benefit for these cases. To this end, we propose a simple but efficient uncertainty estimation ensemble strategy. Specifically, the uncertainty estimation can be calculated as:

$$HR = \frac{N\left(\left(\sum_{i=0}^{C} P_i\right) = \left[\frac{C}{2}\right]\right)}{N\left(\left(\sum_{i=0}^{C} P_i\right) \neq 0\right)} \tag{4}$$

where C and P_i denote the number of model and the result of i_{th} model. $N(\cdot)$ represents the function of finding the number of pixels that meet the condition. We argue that those cases with low HR values tend to have high uncertainty, and tend to have small intracranial hemorrhage areas with blurred boundaries, making it difficult to be effectively segmented by the network. Therefore, we use voting method to do ensemble. For those cases with low HR values, we use fewer votes to get the final result.

$$P_n = \begin{cases} 1, & (HR < th \cup n > N_1) \cap (HR > th \cup n > N_2) \\ 0, & \text{others} \end{cases} \tag{5}$$

where P_n is the final prediction in the pixel n. th indicates half ratio threshold, n is the number of model voting in the pixel and N_1, N_2 are two threshold determine the final results. In our experiments, we use 15 models to ensemble and set $th = 0.5$, $N_1 = 1$ and $N_2 = 7$.

4 Results

Table 1 showed the Dice, hausdorff distance and relative volume difference computed by the competition organizers and displayed in the public leaderboard. Models using $Lwce$ showed better performance and our uncertainty estimation increases the performance slightly.

5 Conclusion

In this work, we described our methodology for intracranial hemorrhage segmentation. We pointed out some challenges for this task and extended nn-UNet by

Table 1. INSTANCE 2022 validation dataset results. Mean Dice, Hausdorff distance and Relative Volume Difference are shown in the table.

Methods	Dice↑	HD↓	RVD↓
2D nnU-Net	0.7066±0.31	null	0.2977±0.32
2D nnU-Net with L_{wce}	0.7257±0.28	44.49±22.75	0.2890±0.26
3D nnU-Net with L_{wce}	0.7393±0.25	31.76±31.36	0.2495±0.25
Simple ensemble 15 models	0.7503±0.24	29.07±26.12	0.2301±0.22
Uncertainty estimation ensemble	**0.7612±0.23**	**29.02±26.34**	**0.2056±0.20**

modifying its loss function and use uncertainty estimation ensemble strategy to address the challenges. Experimental results with unseen validation data from the leaderboard showed the effects of our methodology.

References

1. Avants, B.B., Tustison, N.J., Wu, J., Cook, P.A., Gee, J.C.: An open source multivariate framework for n-tissue segmentation with evaluation on public data. Neuroinformatics **9**(4), 381–400 (2011) 2
2. Bhadauria, H.S., Dewal, M.: Intracranial hemorrhage detection using spatial fuzzy c-mean and region-based active contour on brain ct imaging. Signal, Image and Video Processing **8**(2), 357–364 (2014) 2
3. Goldstein, J.N., Gilson, A.J.: Critical care management of acute intracerebral hemorrhage. Current treatment options in neurology **13**(2), 204–216 (2011) 1
4. Hakimi, R., Garg, A.: Imaging of hemorrhagic stroke. CONTINUUM: Lifelong Learning in Neurology **22**(5), 1424–1450 (2016) 2
5. Hemphill III, J.C., Greenberg, S.M., Anderson, C.S., Becker, K., Bendok, B.R., Cushman, M., Fung, G.L., Goldstein, J.N., Macdonald, R.L., Mitchell, P.H., et al.: Guidelines for the management of spontaneous intracerebral hemorrhage: a guideline for healthcare professionals from the american heart association/american stroke association. Stroke **46**(7), 2032–2060 (2015) 1
6. Ironside, N., Chen, C.J., Mutasa, S., Sim, J.L., Marfatia, S., Roh, D., Ding, D., Mayer, S.A., Lignelli, A., Connolly, E.S.: Fully automated segmentation algorithm for hematoma volumetric analysis in spontaneous intracerebral hemorrhage. Stroke **50**(12), 3416–3423 (2019) 2
7. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021) 1, 2
8. Islam, M., Sanghani, P., See, A.A.Q., James, M.L., King, N.K.K., Ren, H.: Ichnet: intracerebral hemorrhage (ich) segmentation using deep learning. In: International MICCAI Brainlesion Workshop. pp. 456–463. Springer (2018) 1
9. Kwon, D., Ahn, J., Kim, J., Choi, I., Jeong, S., Lee, Y.S., Park, J., Lee, M.: Siamese u-net with healthy template for accurate segmentation of intracranial hemorrhage. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 848–855. Springer (2019) 2
10. 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 **26**(3), 1140–1151 (2021) 1, 2

11. Liu, X., Song, L., Liu, S., Zhang, Y.: A review of deep-learning-based medical image segmentation methods. Sustainability **13**(3), 1224 (2021) 2
12. Maas, A.L., Hannun, A.Y., Ng, A.Y., et al.: Rectifier nonlinearities improve neural network acoustic models. In: Proc. icml. vol. 30, p. 3. Citeseer (2013) 3.2
13. Muschelli, J., Sweeney, E.M., Ullman, N.L., Vespa, P., Hanley, D.F., Crainiceanu, C.M.: Pitchperfect: Primary intracranial hemorrhage probability estimation using random forests on ct. NeuroImage: Clinical **14**, 379–390 (2017) 2
14. Organization, W.H.: World health statistics 2015. World Health Organization (2015) 2
15. Prakash, K., Zhou, S., Morgan, T.C., Hanley, D.F., Nowinski, W.L.: Segmentation and quantification of intra-ventricular/cerebral hemorrhage in ct scans by modified distance regularized level set evolution technique. International journal of computer assisted radiology and surgery **7**(5), 785–798 (2012) 1
16. Roh, D., Sun, C.H., Murthy, S., Elkind, M.S., Bruce, S.S., Melmed, K., Ironside, N., Boehme, A., Doyle, K., Woo, D., et al.: Hematoma expansion differences in lobar and deep primary intracerebral hemorrhage. Neurocritical care **31**(1), 40–45 (2019) 1
17. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015) 2
18. Sacco, S., Marini, C., Toni, D., Olivieri, L., Carolei, A.: Incidence and 10-year survival of intracerebral hemorrhage in a population-based registry. Stroke **40**(2), 394–399 (2009) 1
19. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Instance normalization: The missing ingredient for fast stylization. arXiv preprint arXiv:1607.08022 (2016) 3.2
20. Van Asch, C.J., Luitse, M.J., Rinkel, G.J., van der Tweel, I., Algra, A., Klijn, C.J.: 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 **9**(2), 167–176 (2010) 1
21. Wang, J., Zhu, H., Wang, S.H., Zhang, Y.D.: A review of deep learning on medical image analysis. Mobile Networks and Applications **26**(1), 351–380 (2021) 2
22. Webb, A.J., Ullman, N.L., Morgan, T.C., Muschelli, J., Kornbluth, J., Awad, I.A., Mayo, S., Rosenblum, M., Ziai, W., Zuccarello, M., et al.: Accuracy of the abc/2 score for intracerebral hemorrhage: systematic review and analysis of mistie, clear-ivh, and clear iii. Stroke **46**(9), 2470–2476 (2015) 1