Deeply supervised network for white matter hyperintensities segmentation with transfer learning

Yilei Wu *[1] and Juan Helen Zhou †[1]

[1]National University of Singapore

1 Data Preprocessing

The training dataset consist of 60 subjects from three sites (Utrecht, Singapore, Amsterdam). To train our 3D model, our data preparation included three steps: interpolation, center crop and normalization. For each FLAIR image, we did the interpolation along the z-axis by repeating axial slices three times. Then we took the center crop to make all training image into the shape of 160*192*160. At last, z-score normalization was applied to the whole training dataset.

2 Method

2.1 Model

Our model referred to the 3D U-Net [1] as the base model. A network trained with UKBiobank brain age prediction task [3] is used as the pretrained encoder (Figure 1). We implemented deep supervision [2] by connecting output layers at decoders to generate output with different resolutions. Specifically, our model added output layers at the third and fourth upsampling blocks, thus generating outputs with $\frac{1}{4}$ and $\frac{1}{2}$ of the spatial resolution of the final output respectively. We then downsampled the corresponding labels to calculate the auxiliary loss at each resolution.

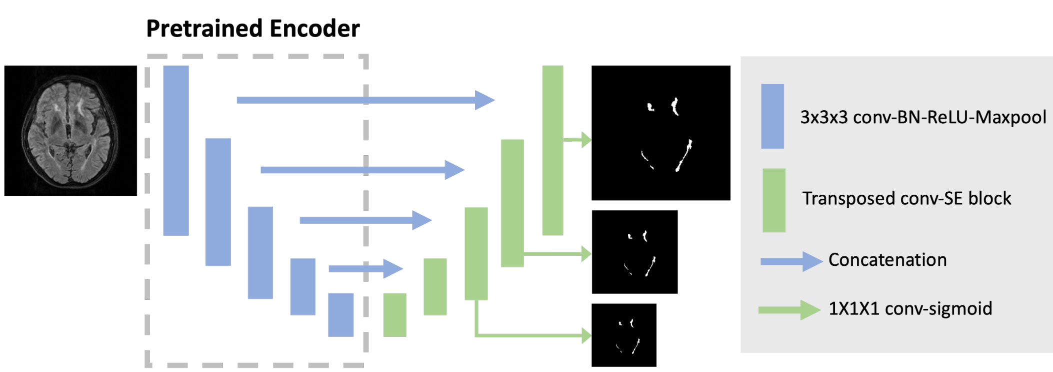

Figure 1: Proposed network structure

2.2 Data augmentation

We applied on-the-fly data augmentation with brightness change (0.2), contrast change (0.25), mirroring (0.5), resampling (0.25), and rotation (0.5); numbers in parentheses refer to the probability per

*yilei.wu@u.nus.edu

†helen.zhou@nus.edu.sg

sample of being transformed during the training.

3 Postprocessing

We generated ensemble result of 10 models from the cross validation. The threshold of 0.1 was applied to the segmentation result to generate the binary mask (WMH v.s non-WMH). Then we reversed the data preprocessing process as mentioned in section 1. To reverse the interpolation, we took the agreement for every three axial slices by calculating the average of them.

References

[1] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *MICCAI*, pages 424–432. Springer, 2016.

[2] Qi Dou, Hao Chen, Yueming Jin, Lequan Yu, Jing Qin, and Pheng-Ann Heng. 3d deeply supervised network for automatic liver segmentation from ct volumes. In *MICCAI*, pages 149–157. Springer, 2016.

[3] Han Peng, Weikang Gong, Christian F Beckmann, Andrea Vedaldi, and Stephen M Smith. Accurate brain age prediction with lightweight deep neural networks. *Medical image analysis*, 68:101871, 2021.