3D Anisotropic Pyramidal Convolution Network for Brain Tissue Segmentation

Yuzhou Zhuang[1], Daiyang Liu [1], Tengying Liu [1]

[1]Huazhong University of Science and Technology, Wuhan, China

zhuangyuzhouhust@gmail.com, daiyangliu562@gmail.com, liutengyingsgg@gmail.com,

In order to overcome the difficulties caused by the anisotropic spatial resolution of medical images and exploit the complementary information in brain multimodal MRI, we propose a novel 3D fully convolutional network called 3D Anisotropic Pyramidal Convolution Network, which consists of an anisotropic pyramidal convolutional encoder and a multi-scale feature fusion decoder. To address the problem of memory footprint, we also use more memory-efficient residual architecture. Moreover, we design spatial attention module to make the network focus more on the foreground region. A paper describing the methods is currently in preparation and will be made public as soon as possible. A short description of the method is given in the following.

The proposed method is fully automatic, and we used all three modalities (i.e., T1, T1 IR and FLAIR) for training and inferencing. In the training phase, we input randomly cropped sub-volumes of the training scans of size $32 \times 196 \times 196 \times N$ for training our proposed fully convolutional network, where N is the number of MRI modalities ($N = 3$ for the T1, T1-IR and T2-FLAIR). Our model is trained for 100 epochs with initial learning rate 0.0001, and the learning rate is halved every 15 epochs. In the inferencing phase, we apply the trained model on the scans to be segmented using a sliding window with overlap of 8.

The algorithm is designed to segment brain tissue in multimodal MRI, and it is aimed to be applicable for most people (e.g. scans of healthy volunteers and patients with specific pathology).

We only use the training data provided for the MRBrainS13 challenge to train model without using other datasets, and we don't use transfer learning to initialize the weights of the proposed network. We also use classical data augmentation via horizontal/vertical flipping, rotation of the original MRI.

Our algorithm only uses gray matter, white matter and cerebrospinal fluid labels, and we merge labels 1 and 2, 3 and 4, and 5 and 6 in training process and inferencing process. Besides, our algorithm hasn't been tested on other databases.

The average runtime of our algorithm in inferencing phrase is about 45 seconds averaged on all the 15 test images using one NVIDIA GeForce GTX 1080Ti GPU on Ubuntu 16.04.