An Efficient Brain-Tissue Segmentation Algorithm based on 3D UNet and Uncertain Focal Loss Function

Shuai Chen[1], Hua Ma[1], Chengjia Wang[2], and Marleen de Bruijne[1,3]

[1] Biomedical Imaging Group Rotterdam, Departmenst of Medical Informatics, Radiology and Nuclear Medicine, Erasmus MC, Rotterdam, The Netherlands
[2] BHF Centre for Cardiovascular Science , University of Edinburgh, The United Kingdom
[3] Department of Computer Science, University of Copenhagen, Denmark

1 Data

7 MRI datasets acquired on a 3T scanner at the UMC Utrecht (the Netherlands) were used as the training set of our method. For each of the subject, fully annotated multi-sequence (T1-weighted, T1-weighted inversion recovery and T2-FLAIR) scans are available. There are another 23 unseen datasets as testing sets of the challenge, which are not accessible. The 30 subjects include patients with diabetes, dementia and Alzheimer's disease, and subjects with increased cardiovascular risk, with varying degrees of atrophy and white matter lesions (age>50).

2 Preprocessing

2.1 Patches extraction

Due to the limitation of GPU memory, we use patches of the size 200*200*16 voxels as the input to our networks. The patches are extracted from the 3D region of interest (cropping the background outside the center of each slice in axial direction, from the size 240*240*16 into 200*200*16) of the 3D MR scans, which contains most of the whole brain. There is 90% overlap between neighbor patches in training.

2.2 Data augmentation

Several 3D data augmentation strategies were applied on the training patches, including 3D rotation, shifting, as well as flipping in 3 directions (axial, sagittal, coronal). For detailed parameters, rotation in 3 directions are [10, 5, 5] degrees, shifting range in 3 directions are [24, 35, 7] voxels.

3 Method

3.1 3D UNet

The main CNN architecture of our method is based on the UNet[1]. The illustration of the network and parameters is in Figure 1.

3.2 Uncertain focal loss function

We used the *uncertain focal categorical cross-entropy* as loss function. The modification we made to the regular categorical cross-entropy loss function is adding a focal term [2] to address the class imbalance problem and an uncertainty term to that prevents the model from forcing the output probabilities to be close to 0 or 1 even for ambiguous areas. For example, in 3-class segmentation, the loss function is:

$$loss_{class\ i} = -\left(1-z_i\right)^{\gamma}(1-\varepsilon)\log\left(e^{z_i}/Z\right) - \left(1-z_i\right)^{\gamma}\varepsilon\sum_{i=1}^{n}\frac{1}{3}\log\left(e^{z_i}/Z\right),\ Z = e^{z_1}+e^{z_2}+e^{z_3}$$

where z_i is the output of the last CNN layer and e^{z_i} is the form after *softmax* activation function. The parameter γ could be set between (0, 5) according to the focal loss paper [2]. Here we use the optimal value $\gamma = 2$ as described in [2]. The parameter ε determines how much "uncertainty" the solution should have. The value of ε we set depends on the quality of the ground truth. In our experiments $\varepsilon =$ 0.01. To minimise this cost function we used the ADAM optimiser [24] with a learning rate of 1.0e-5.

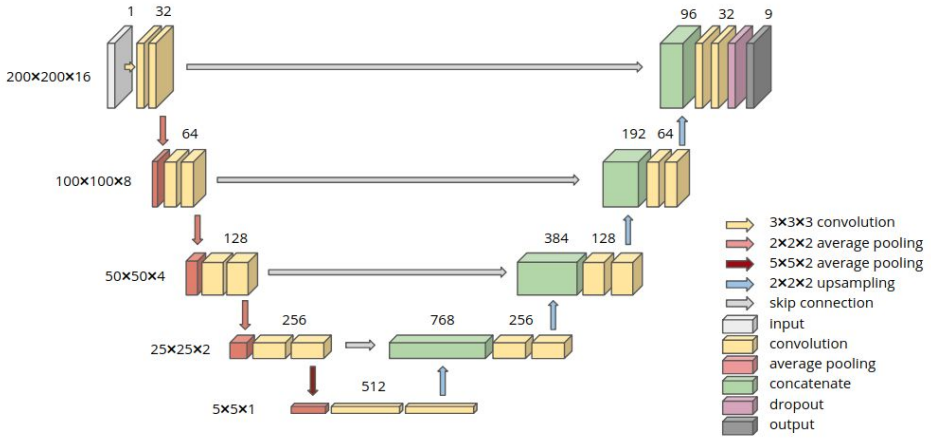

Figure 1: 3D UNet

References

1. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.
2. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). Focal Loss for Dense Object Detection. Proceedings of the IEEE International Conference on Computer Vision, 2017–Octob, 2999–3007.