Deep Attentive Network with Mixed Kernels Convolution for Brain Segmentation

Haoran Dou[1 2], Davood Karimi[2] and Ali Gholipour[2]

[1] School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China
[2] Department of Radiology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA
Haoran.Dou@childrens.harvard.edu

Abstract. we proposed a novel deep attentive neural network for brain segmentation. Our proposed network leverages the attention mechanism with a mixed kernels convolution to capture multi-scale contextual information. This enables our model to learn to accurately segment the fine and complex structure of cortical plate.

Keywords: Attention mechanism· Brain segmentation· Deep learning.

1 Methods

our network works on patches. For each input patch, the network outputs a probability map of the same size in an end-to-end manner. The network first extracts a series of feature maps with different resolutions. The shallow features contain rich boundary detail information with high resolutions while the deep features comprise the high-level semantics. We implement the 3D U-Net[1] as our backbone network. The network down sampling four times in the encoder of the backbone network and the number of features at each stage is 16, 32, 64, 128, 128, respectively.

After obtaining the feature maps with different levels of context information via backbone network, we up sample these features with different resolutions to the same size of input by trilinear interpolation. Then, a convolution with kernel size of 1x1x1 is employed to compress the channel of features to 16. After that, the deep supervision[2] is employed to enhance the back propagation of gradient flows of the whole networks and help network to efficiently learn discriminative represents in the feature maps.

To address the large scale difference in the fetal brain MRI due to the rapid morphological change during fetal brain development. we propose a stage-wise attention module to refine the feature maps with different resolutions. The developed attention module leverages mixed kernels convolution to extract multi-scale information and refine the feature maps in a stage-wise manner.

Finally, we use the deep supervision[2] to strength the characteristics in the feature maps after attentive refinement and combine these attentive features together. After processed by two convolution layers following with a BN and Prelu, the network outputs the final probability map.

Bibliography

[1] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016). 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer.

[2] Xie, S. and Tu, Z. (2015). Holistically-nested edge detection. In *Proceedings of the IEEE international conference on computer vision*, pages 1395–1403.