DESCRIPTION

Yunhe Gao[1]

[1] The Chinese University of Hong Kong, Sha Tin, Hong Kong
yhgao@link.cuhk.edu.hk

Abstract. Intervertebral discs are joints that lie between vertebras in the spinal column, which absorb shock between vertebrae during activity. There is a strong correlation between lower back pain and degeneration of intervertebral discs, which may have a great impact on people's normal life, especially for people who sit for a long time at work. The precise segmentation of the intervertebral disc is of great significance for the diagnosis of disc degeneration. Currently clinical practice usually manually annotates the volumetric data, which is time-consuming and tedious. In this challenge, we developed a fully automated algorithm based on fully convolutional neural network architecture that can accurately segments and locates seven intervertebral discs.

Keywords: Intervertebral Discs, Segmentation, Localization, Deep Learning

1 Method

The entire pipeline mainly consists of two parts. Disc segmentation and localization. The segmentation network takes the raw four modality images as input, and outputs the binarized segmentation results of intervertebral discs. Since the original MR image includes seven target intervertebral discs and several thoracic discs with similar appearance, the segmentation network cannot distinguish the difference. Therefore, we further used a localization network to distinguish the seven target discs from the other thoracic discs to achieve the goal of fully automatic

For the segmentation networks, we used a 2D fully convolutional neural network [2]. Although many studies have shown that the 3D methods have stronger capability to learn 3D discriminative features from volumetric images, we observed that the label are annotated in 2D slices, and many discontinuities appear in coronal plane. We think it might bring problems to the performance of 3D networks. Moreover, 3D method has a larger demand for data, however, the training set has only 16 samples. Therefore, we think the 2D network is more suitable for this task. Out segmentation network is similar to UNet [1], using the Encoder-Decoder structure, but we only performed two down-samples. In order to ensure a large enough receptive field, we used dilated convolution.

The localization network that distinguishes 7 intervertebral discs from other thoracic discs is a 3D network, which is trained for multi-class classification, with the raw 4 modality images as input, and outputs the labels of 7 lumbar discs as 1-7 from bottom to up. To further enhance the robustness of the results, we used a registration method combined with the results of the network to ensure accurate separation of 7 discs.

It should be noted that we used additional process to further boost the performance. It is a natural assumption that only the spine part of the entire input image is useful for the segmentation of the intervertebral disc, the part outside the spine is only useless background that may confuse the network from accurate segmentation. Therefore, we first train a network to predict the spine's mask, ignoring the loss of the area outside the mask while training the segmentation network, so that the segmentation network can focus within the spine area. The ground true label of this network is obtained from the convex hull of the label of discs after some dilation.

In this submitted docker, we used several modified models based on UNet for ensemble to further improve performance and robustness.

References

1. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.
2. Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.