Vertebral Column Segmentation Using Single-staged CNNs

Sevde Aydogdu[1], Danail Stoyanov[1], Deepak M. Kalaskar[2], and Evangelos Mazomenos[1]

[1] Wellcome / EPSRC Centre for Interventional and Surgical Sciences, University College London, London, UK
sevde.aydogdu.21@ucl.ac.uk
[2] Department of Orthopaedics and Musculoskeletal Science, University College London, London, UK

1 Introduction

Spine segmentation from computed tomography (CT) is a key step in the diagnosis and successive treatment of spinal conditions. Spine segmentation is crucial in clinical practice, allowing medical personnel to evaluate the morphology and identify abnormalities and deformations. Improper segmentation can lead to misdiagnosis and suboptimal treatment strategies. The majority of the state-of-the-art in automated spine segmentation are typically complex pipelines with multiple processing stages [3]. In this study, we evaluate several Convolutional Neural Network (CNN) architectures in a single-stage segmentation pipeline.

2 Methodology

Single-stage CNNs based on U-Net[2], U-Net++[5] and Attention U-Net[4] were used, and evaluated on the Verse20 dataset from MICCAI 2020 challenge[3]. The effect on performance of dense skip connections and attention (AG) components was explored. The CT scans were pre-processed to correct voxel variation and orientation throughout the dataset. The CT scans were split into a 75:15:10 ratio for training, validation, and testing. Training was performed with a batch size of 16 using ADAM optimization, for a total of 100 epochs. A combination of the binary cross-entropy and dice loss was used as the loss function.

3 Results & Discussion

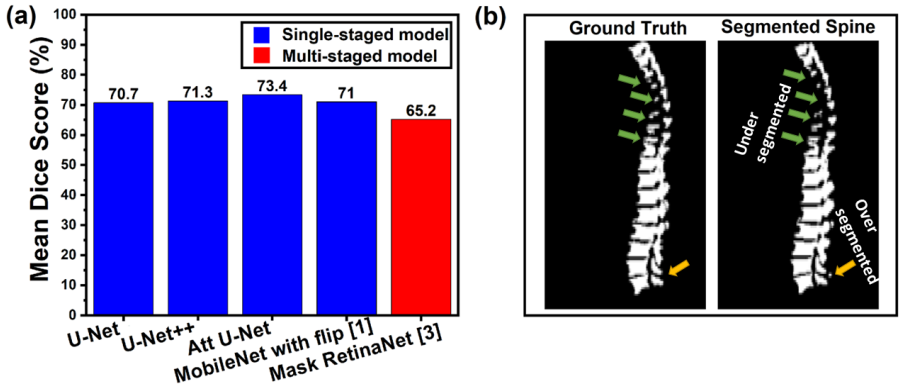

Fig. 1. (a) Comparison segmentation performance for Verse'20, and (b) Qualitive comparison of Att U-Net by visulazition

Segmentation performance is evaluated with the dice score (DSC) and results are shown in Fig. 1. Attention U-Net achieved the highest DSC, increased by 2.7% over the standard U-Net. U-Net++ showed slightly better performance compared to U-Net thanks to recovering more spatial information via dense skip connections. Our preliminary findings show that AG can improve the learning process of the segmentation task, enhancing the effectiveness of skip connections, since it provides paying attention to the target region (spine). Taking advantage of AG, improves DSC accuracy by approximately 2.4% and 8.2% compared to [1] and [3], respectively. AG offers the flexibility of directing the CNN to focus exclusively on the regions related to the segmentation task. Overall, Attention U-net appears to be capable of finer segmentation (Fig. 1.), however, there are some under- and over-segmented details regarding shape and boundary information. We conclude that by utilising AG, spine segmentation in CT can be achieved in a single-stage CNN pipeline. In our future work, we will focus on a single-stage, attention-based architecture where AGs are integrated in the feature extraction part of the CNN, to refine the segmentation task.

References

1. Hempe, H., Heinrich, M.P.: Towards single-stage prediction of vertebrae segmentation for opportunistic screening of 3d ct scans.(2021)
2. Ronneberger et al.: U-net: Convolutional networks for biomedical image segmentation.MICCAI,(2015)
3. Sekuboyina, A. et al.: Verse: a vertebrae labelling and segmentation benchmark for multi-detector ct images.MIA,(2021)
4. Vaswani, A. et al.: Attention is all you need.ANIPS,(2017)
5. Zhou, Z. et al.: Unet++: A nested u-net architecture for medical image segmentation.DLMIA, and ML-CDS,(2018)