 DOI: 10.3217/978-3-85125-603-1-02

Multi-Label Whole Heart Segmentation Using Anatomical Label Configurations and CNNs

Christian Payer[1], Darko Štern[2], Horst Bischof[1], Martin Urschler[2]

I. Extended Abstract

The accurate analysis of the whole heart substructures, i.e., left and right ventricle, left and right atrium, myocardium, pulmonary artery and the aorta, is highly relevant for cardiovascular applications. Therefore, automatic segmentation of these substructures from CT or MRI volumes is an important topic in medical image analysis [4]. Challenges for segmenting the heart substructures are their large anatomical variability in shape among subjects, the potential indistinctive boundaries between substructures and, especially for MRI data, artifacts and intensity inhomogeneities resulting from the acquisition process. To objectively compare and analyze whole heart substructure segmentation approaches, efforts like the MICCAI 2017 Multi-Modality Whole Heart Segmentation (MM-WHS) challenge are necessary and important for potential future application of semi-automated and fully automatic methods in clinical practice. We participated in the MM-WHS challenge, where we proposed a deep learning framework for fully automatic multi-label segmentation [2]. Evaluated on the MM-WHS challenge test data, we rank first for CT and second for MRI with a whole heart segmentation Dice score of 90.8% and 87%, respectively, leading to an overall first ranking among all participants.

Our proposed method [2] performs fully automatic multi-label whole heart segmentation with CNNs using volumetric kernels. Due to the extensive memory and runtime requirements of volumetric CNNs, we use a pipeline of two CNNs that first *localizes* the heart in lower resolution volumes to crop a standardized region around the heart, followed by obtaining the final *segmentation* in higher resolution within this region (see Fig. 1).

The *localization* CNN based on the U-Net [3] performs landmark localization using heatmap regression [1] to predict the approximate center of the bounding box around all heart substructures. Then, we crop a region with fixed physical size around the predicted center, ensuring that the region encloses all segmentation labels. Within the cropped region, the multi-label *segmentation* CNN predicts the heart substructure labels of each voxel. For this task, we use an adaptation of the fully convolutional end-to-end trained SpatialConfiguration-Net (SCN) from [1]. The main idea of the three-component SCN is to learn from relative positions among structures to focus on anatomically feasible configurations as seen in the training data.

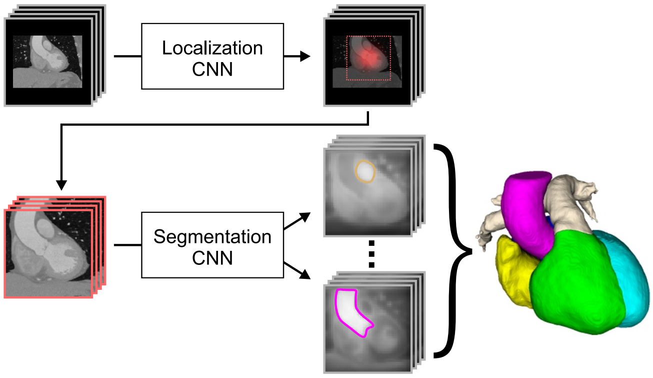

Fig. 1. Overview of the fully automatic two-step multi-label segmentation pipeline. The first CNN uses a low resolution volume as input to localize the center of the bounding box around all heart substructures. The second CNN crops a region around this center and performs the multi-label segmentation. The figure is adapted from [2].

In the first component of the SCN, a U-Net-like architecture [3] generates the first intermediate label predictions, corresponding to a voxel-wise probability of all labels. Then, the second component models the spatial configuration among labels, by using consecutive convolution layers to transform these probabilities to positions of other labels, generating the second intermediate label predictions. Finally, the third component multiplies both intermediate predictions, which results in the combined label predictions. Note that only when trained in an end-to-end manner, this last multiplication ensures that the first and second network component perform as expected. Without any further postprocessing, choosing the maximum value among the label predictions for each voxel leads to the final multi-label segmentation.

*This work was supported by the Austrian Science Fund (FWF): P 28078-N33.

[1]Christian Payer and Horst Bischof are with the Institute of Computer Graphics and Vision, Graz University of Technology, Austria christian.payer@icg.tugraz.at

[2]Darko Štern and Martin Urschler are with the Ludwig Boltzmann Institute for Clinical Forensic Imaging, Graz, Austria martin.urschler@cfi.lbg.ac.at

References

[1] C. Payer, D. Štern, H. Bischof, and M. Urschler, "Regressing Heatmaps for Multiple Landmark Localization Using CNNs," in *Proc. Med. Image Comput. Comput. Interv.* Springer, 2016, pp. 230–238.

[2] C. Payer, D. Štern, H. Bischof, and M. Urschler, "Multi-label Whole Heart Segmentation Using CNNs and Anatomical Label Configurations," in *Stat. Atlases Comput. Model. Hear. ACDC MMWHS Challenges. STACOM 2017.* Springer, 2018, pp. 190–198.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Proc. Med. Image Comput. Comput. Interv.* Springer, 2015, pp. 234–241.

[4] X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI," *Med. Image Anal.*, vol. 31, pp. 77–87, 2016.