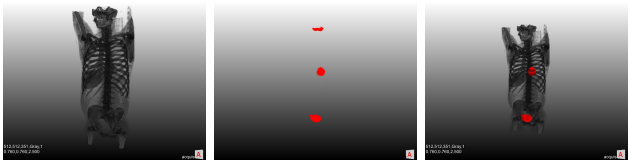
Learning from the Truth: Fully Automatic Ground Truth Generation for Training of Medical Deep Learning Networks*

Christina Gsaxner[1,2,3], Peter M. Roth[1], Jürgen Wallner[2,3] and Jan Egger[1,2,3]

I. PROBLEM STATEMENT AND MOTIVATION

Automatic medical image analysis has become an invaluable tool in the different treatment stages of diseases. Especially medical image segmentation plays a vital role, since segmentation is often the initial step in an image analysis pipeline. Convolutional neural networks (CNNs) have rapidly become a state of the art method for many medical image analysis tasks, such as segmentation. However, in the medical domain, the use of CNNs is limited by a major bottleneck: the lack of training data sets for supervised learning. Although millions of medical images have been collected in clinical routine, relevant annotations for those images are hard to acquire. Generally, annotations are created (semi-)manually by experts on a slice-by-slice basis, which is time consuming and tedious. Therefore, available annotated data sets are often too small for deep learning techniques.



(a) CT data (b) PET data (c) PET/CT

Fig. 1: 3D image data obtained from CT, PET, and combined PET/CT. In CT data (a), contrast for soft tissue is poor. PET data in (b) shows metabolical active regions. A PET/CT scan (c) allows to properly assign active regions anatomically.

II. METHOD OVERVIEW

To overcome these problems, we proposed a novel method to automatically generate ground truth annotations by exploiting positron emission tomography (PET) data acquired simultaneously with computed tomography (CT) scans in combined PET/CT systems [3], [4]. PET/CT scanning combines functional information from PET with anatomical information from CT. Soft tissue, which exhibits limited contrast in CT, shows up distinctively in a PET scan if it is metabolically active (see Figure 1). However, PET scanning increases radiation exposure for the patient and is not as widely available as CT, making approaches which detect significant structures in CT scans alone compelling. We utilized the high contrast in PET scans to extract ground truth segmentations for corresponding structures of interest in CT data, enabling automatic detection of these structures in CT alone by training CNNs with the generated data. As a structure of interest we chose the urinary bladder, since the radio-tracer used for PET imaging always accumulates in it. The ground truth is acquired fully automatically from PET by a thresholding algorithm. Furthermore, affine transformations and noise are applied to the generated data for data augmentation [2]. Using these data, we trained and tested different CNN architectures for image segmentation, which are based on fully convolutional networks [5] and Deeplab [1].

III. RESULTS AND DISCUSSION

Qualitative segmentation results predicted with our best performing architecture are shown in Figure 2. Quantitatively, we achieve a maximal mean Dice coefficient of 81.9%. This results are very satisfactory, considering that no manually annotated training data was used in our studies. Our approach presents a promising tool for automatic CT analysis and can be generalized to all applications of PET/CT. In particular, in future research we aim to extend our method to tumor detection.

Fig. 2: Qualitative segmentation results. The prediction is shown in red, while the ground truth is outlined in green.

*Supported by FWF KLI 678-B31 (enFaced), COMET K-Project 871132 (CAMed) and the TU Graz Lead Project (Aortic Dissection).

[1]Institute of Computer Graphics and Vision, TU Graz, Austria
[2]Computer Algorithms for Medicine Laboratory, Graz, Austria
[3]Department of Oral and Maxillofacial Surgery, Medical University of Graz, Austria

REFERENCES

[1] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE PAMI*, 40(4):834–848, 2018.

[2] C. Gsaxner, B. Pfarrkirchner, L. Lindner, N. Jakse, J. Wallner, D. Schmalstieg, and J. Egger. Exploit 18F-FDG enhanced urinary bladder in PET data for deep learning ground truth generation in CT scans. In *Proc. SPIE Medical Imaging*, 2018.

[3] C. Gsaxner, B. Pfarrkirchner, L. Lindner, A. Pepe, J. Wallner, P. M. Roth, and J. Egger. PET-Train: Automatic Ground Truth Generation from PET Acquisitions for Urinary Bladder Segmentation in CT Images using Deep Learning. In *Proc. BMEiCON*, 2018.

[4] C. Gsaxner, P. M. Roth, J. Wallner, and J. Egger. Exploit fully automatic low-level segmented PET data for training high-level deep learning algorithms for the corresponding CT data. *PLOS ONE*, 14(3):1–20, 2019.

[5] J. Long, E. Shelhamer, and T. Darrell. Fully Convolutional Networks for Semantic Segmentation. In *Proc. CVPR*, 2015.