Comparison of Automatic Segmentation Methods for Total Body PET/CT Imaging

Anting Li[1,2], Jarmo Teuho[1] , Vilma Itkonen[1] , Maria K. Jaakkola[1] , Riku Klén[1]
Turku PET Centre, University of Turku and Turku University Hospital, Turku, Finland
Faculty of Medicine, University of Turku, Turku, Finland

INTRODUCTION

Segmenting regions of interest from total body Positron Emission Tomography/Computed Tomography (PET/CT) images is time-consuming and susceptible to variability between operators. Especially when multiple organs need to be delineated and analyzed. Automatic segmenting tools have been developed to address these challenges, for example, machine learning approaches, which allow fast processing of extensive datasets with good accuracy. In this study, we exploit two CT-segmenting tools, TotalSegmentator [1] and MIWBAS [2], and evaluate their outputs at the practical level and explore their potential applications in total body PET/CT Imaging.

MATERIAL & METHODS

Dataset

All images were obtained from the Biograph Vision Quadra PET/CT system. This scanner provides an axial field of view (AFOV) of 106 cm [3]. Our test data includes 30 total body (from the top of the head to the thigs) CT images scanned with the Biograph Vision Quadra total body PET/CT system.

Algorithms

Two algorithms were evaluated, MIWBAS version 1.0 and TotalSegmentator version 1.5.6. MIWBAS is introduced by Siemens Healthineers. It is based on deep learning, but no details about the model are publicly available. TotalSegmentator is an independent open-source software package based on the U-Net framework. TotalSegmentator tends to use more detailed subsegments whereas MIWBAS segments full organs as single masks.

Overlapping Organs

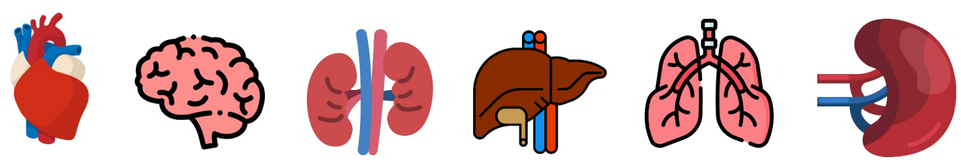

Fig. 1. Overlapping organs include aorta, heart, brain, kidneys, liver, lungs, and spleen. Icons were downloaded from Flaticon. https://www.flaticon.com/

Evaluation

To compare differences between organs segmented by the two methods, we use the Jaccard index as a metric. Jaccard index is defined as the ratio of the size of the overlapping part (number of voxels) and the size of their union (number of voxels).

RESULTS

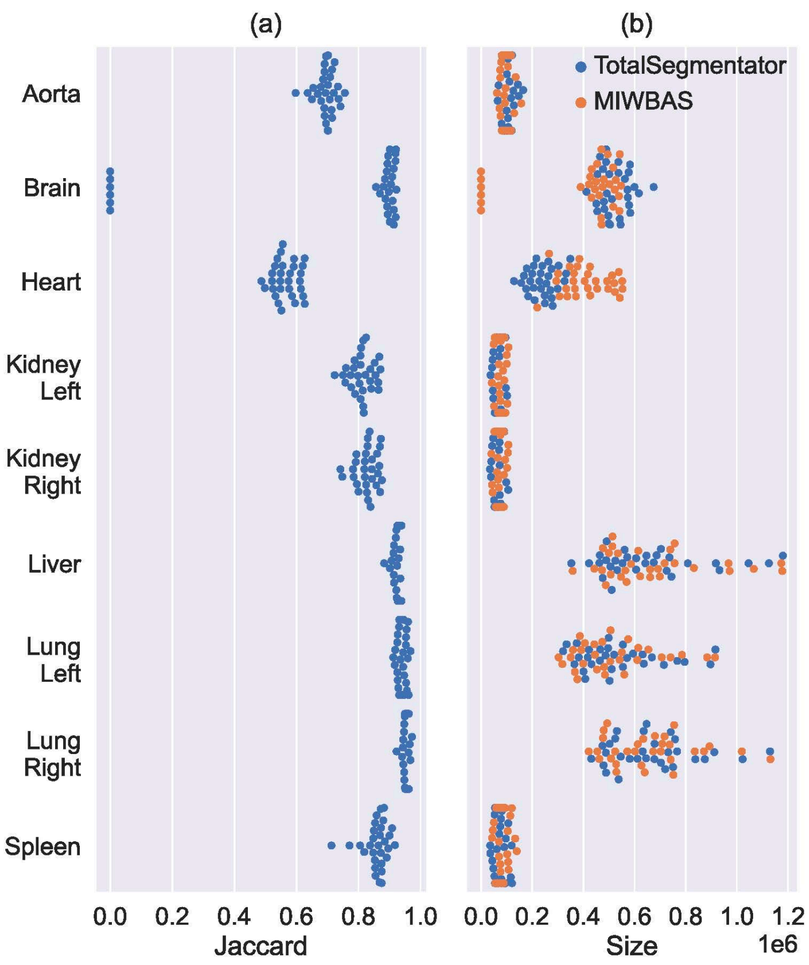

Fig. 2. (a) Jaccard indices describing similarity of TotalSegmentator and MIWBAS segmentation in different organs, and (b) sizes (number of voxels) of different segments by TotalSegmentator (blue) and MIWBAS (orange).

The brain, the lungs, and the liver were the least ambiguous organs to segment and the overlap between the methods was high, resulting in the Jaccard index ≥ 0.9 in most cases (Fig 2a). Notably, in 6 out of the 30 test images MIWBAS did not segment the brain for unknown reason causing the Jaccard index to become zero. Heart had the lowest agreement between the two methods with the mean Jaccard of 0.566. In most organs there were no systematic differences in their segment sizes defined by TotalSegmentator and MIWBAS (Fig 2b). However, MIWBAS segments the heart systematically bigger than TotalSegmentator.

Our preliminary results show a systematic difference between the aorta and heart segmentations of TotalSegmentator and MIWBAS (Figure 3). Typically MIWBAS defines the heart to be bigger than TotalSegmentator, while TotalSegmentator classifies some areas as part of the aorta. In some cases also inner parts of kidneys are systematically segmented as kidneys in MIWBAS, but excluded from the kidney segments defined by TotalSegmentator.

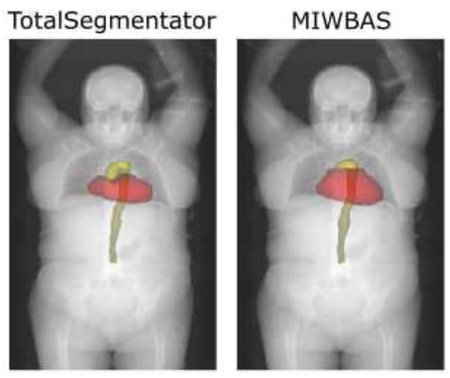

Fig. 3. Typical difference in the segmentation of the aorta (yellow) and the heart (red) performed by TotalSegmentator and MIWBAS.

DISCUSSION AND CONCLUSION

Our preliminary results show that for segmentation of the major organs, both methods are very comparable, except for regions in the brain, aorta, kidneys, and the heart. Comparing the methods is currently challenging due to the different working principles of both tools. Whereas MIWBAS segmentations include complete organs, TotalSegmentator includes several sub-regions with finer detail. To perform the comparison, we combined these sub-sections which might result in loss of small details in the process. This is a preliminary study including only two methods and limited test data. We plan to expand the study by including a similar recently published method MOOSE[4] into the comparison and utilizing more diverse datasets. Furthermore, for absolute ground truth evaluation, manual delineations need to be included.

REFERENCE

[1] J. Wasserthal et al. “TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images,” Radiology: Artificial Intelligence, vol. 5, no. 5, 2023.
[2] R. Seifert et al. “Semiautomatically Quantified Tumor Volume Using 68Ga-PSMA-11 PET as a Biomarker for Survival in Patients with Advanced Prostate Cancer,” Journal of Nuclear Medicine, vol. 61, pp. 1786-1792, 2020.
[3] G. A. Prenosil et al. Performance Characteristics of the Biograph Vision Quadra PET/CT System with a Long Axial Field of View Using the NEMA NU 2-2018 Standard,” Journal of Nuclear Medicine, vol. 63 no. pp. 3 476-3484, 2022
[7] S. Shiyam et al. “Fully-automated, semantic segmentation of whole-body 18F-FDG PET/CT images based on data-centric artificial intelligence,” Journal of Nuclear Medicine, vol 63, no. 12, pp. 1941-1948, 2022