

# Fully automated multi-organ segmentation in CT images via deep neural networks



Yazdan Salimi<sup>1</sup>, Isaac Shiri<sup>1</sup>, Zahra Mansouri<sup>1</sup>, and Habib Zaidi, *IEEE*<sup>1,2,3,4</sup>

<sup>1</sup>Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, Geneva, Switzerland

<sup>2</sup>Geneva University Neurocenter, University of Geneva, 1205 Geneva, Switzerland

<sup>3</sup>Department of Nuclear Medicine and Molecular Imaging, University of Groningen, Groningen, Netherlands

<sup>4</sup>Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark

## INTRODUCTION

- Organ segmentation is a critical task in different medical imaging applications, including radiation therapy treatment planning, image-guided surgeries, personalized dosimetry, and computation quantitative diagnostic evaluations. Delineation of organs by the manual process is labor-intensive and time-consuming, and there are intra-observer variations.
- The low contrast of many tissues in computed tomography (CT) makes manual segmentation more challenging on CT images. There are multiple online available datasets of manual segmentations for organs segmented on CT images. They can be used to train deep neural networks, which have been proven to have a good capacity for the semantic segmentation of healthy organs and pathological lesions. This study aimed to develop a deep network to perform multi-organ segmentation tasks on CT images.

## METHODS

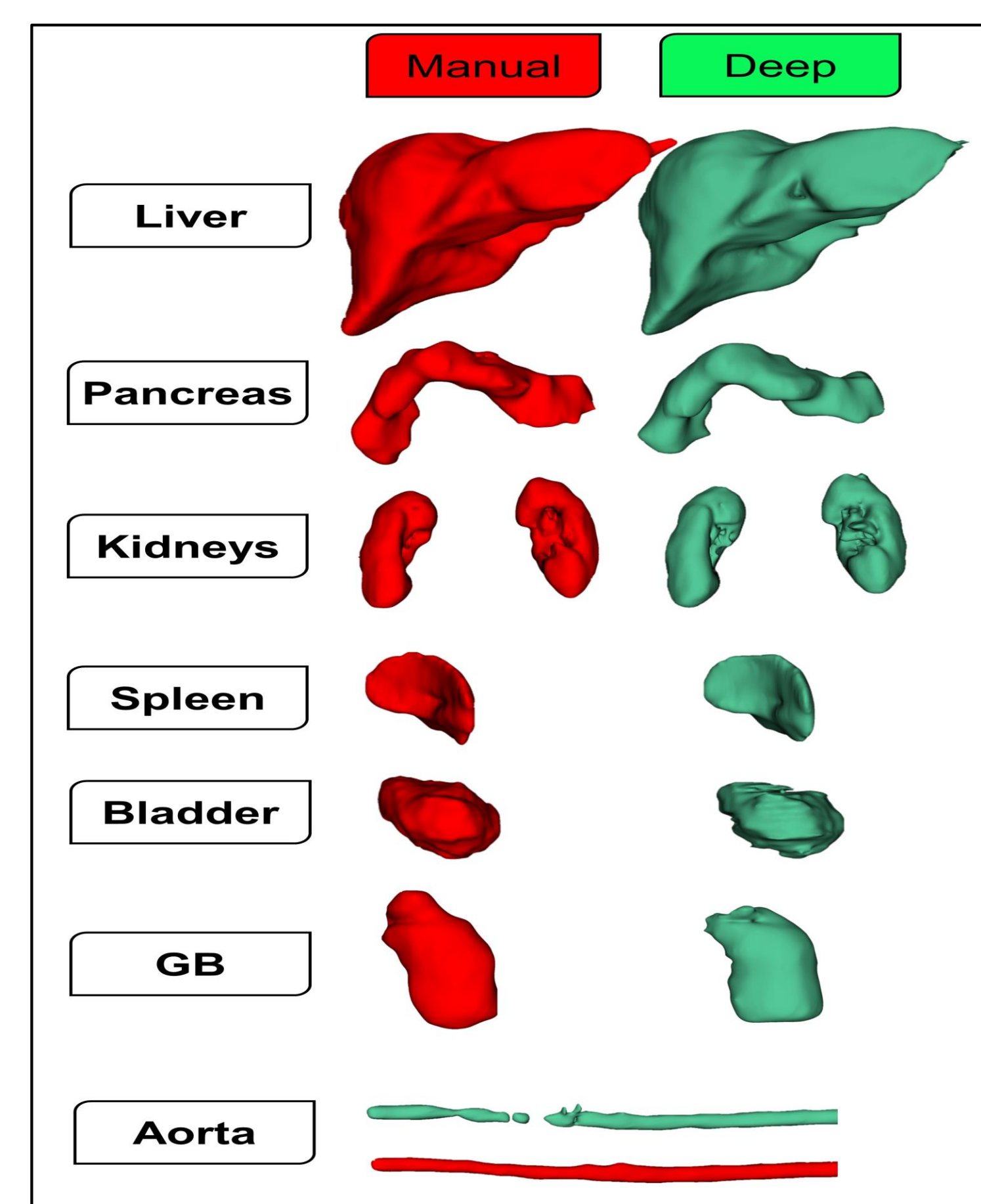
- Multiple datasets, including CT images and segmentations of organs of the Liver, Pancreas, Kidney, Spleen, Bladder, Aorta, and Gallbladder, were collected. Table 1 shows the number of images included in each dataset.
- The body contour was extracted automatically on all the images; after visual assessment of the body contours, all the CT images were cropped to a bounding box containing the patients' body, then clipped between -450 and +450 HU and normalized between zero and one. The resulting image was resized to 200x200 size in the axial plane.
- A residual network, HighRes3DNet, was trained to segment 2D axial images after splitting the data to train and validate (80%) and external test (20%) test sets. For any organ segmentation task, the data splitting was done on each dataset, i.e., there are images from all datasets in train and test sets. The training process was continued for 100 epochs with the Dice loss function and L2 regularization with 0.00009 decay and the initial learning rate equal to 0.0001.
- The network output segmentation was compared to the ground truth masks after post-processing by removing the smaller objects detected in the image. The two deep and ground truth images were compared using the dice similarity coefficient and Jaccard similarity index.

**Table.1.** Number of images from each dataset used in this study.

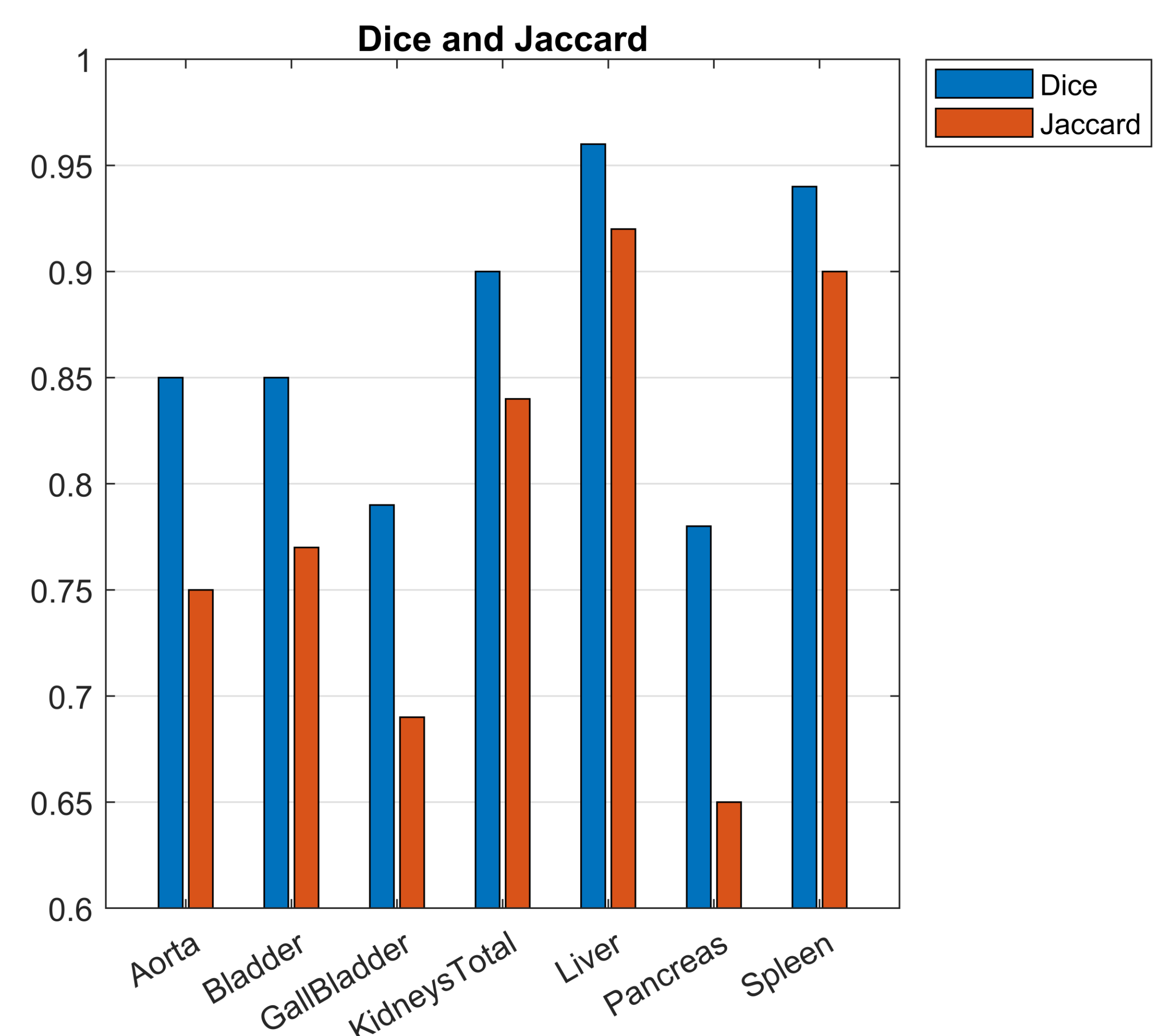
Organ	Dataset Name	Train #	Test #
Liver	Dataset 1	23	7
	Dataset 2	99	28
	All_included	122	35
Pancreas	Dataset 1	25	5
	Dataset 3	62	18
	Dataset 4	222	58
	All_included	309	81
Kidneys	Dataset 5	68	17
	Dataset 6	180	45
	All_included	248	62
Spleen	Dataset 1	24	6
	Dataset 4	34	6
	All_included	58	12
Bladder	Dataset 1	4	1
	Dataset 6	89	18
	All_included	93	19
Aorta	Dataset 1	24	6
	All_included	24	6
Gallbladder	Dataset 1	21	7
	Dataset 7	214	53
	All_included	235	60

## RESULTS

The average dice values were  $0.96 \pm 0.02$ ,  $0.78 \pm 0.2$ ,  $0.90 \pm 0.19$ ,  $0.94 \pm 0.08$ ,  $0.85 \pm 0.16$ ,  $0.85 \pm 0.06$ , and  $0.79 \pm 0.2$  for Liver, Pancreas, Kidney, Spleen, Bladder, Aorta, and Gallbladder, respectively. Also, the Jaccard values were  $0.92 \pm 0.02$ ,  $0.65 \pm 0.20$ ,  $0.84 \pm 0.19$ ,  $0.90 \pm 0.08$ ,  $0.77 \pm 0.18$ ,  $0.75 \pm 0.08$ ,  $0.69 \pm 0.21$  for the mentioned organs, respectively. The best dice were for the liver, followed by the spleen, while the weakest performance was shown in pancreas and gall bladder segmentation.



**Fig.1.** Three-dimensional presentation of ground truth (left, red) segmentations and their corresponding deep (right, green) segmentations. GB: Gall Bladder. All the segments are shown from the anterior point of view except GB and aorta, which are shown from the right lateral.



**Fig.2.** Bar plots of the average Dice and Jaccard metrics achieved in this study.

## CONCLUSION

- In this study, we used different datasets to train a deep residual network to perform fully automatic organ segmentation on CT images.
- The performance of our model was almost comparable to or better than the recently published results by Tang et al., Zhou et al., Huo et al. and Chen et al. and Let al.al for selected organs.
- We used different datasets and average the scores on the external data by combining all available datasets. These data contained different pathologies and combinations of different acquisition parameters with/without contrast enhancement. These results can be developed by newer networks or by using more datasets for training.
- Our developed model can be used as a help in radiation therapy treatment planning or personalized dosimetry on CT scanning by generating patient-specific computational phantoms