Autocontouring of the mouse thorax using deep learning

Justin Malimban[1], Danny Lathouwers[2], Haibin Qian[3], Frank Verhaegen[4], Julia Wiedemann[1], Sytze Brandenburg[1], and Marius Staring[5]

[1]University Medical Center Groningen, Department of Radiation Oncology, The Netherlands
[2]Delft University of Technology, Department of Radiation Science and Technology, The Netherlands
[3]Amsterdam University Medical Centers (location AMC) and Cancer Center Amsterdam, Department of Medical Biology, The Netherlands
[4]Maastricht University Medical Center, Department of Radiation Oncology (MAASTRO), The Netherlands
[5]Leiden University Medical Center, Department of Radiology, The Netherlands

Purpose or Objective

Image-guided small animal irradiations are typically performed in a single session, requiring continuous administration of anesthesia. Prolonged exposure to anesthesia can potentially affect experimental outcomes and thus, a fast preclinical irradiation workflow is desired. Similar to the clinic, delineation of organs remains one of the most time-consuming and labor-intensive stages in the preclinical workflow, and this is amplified by the fact that hundreds of animals are involved in a single study. In this work, we evaluated the accuracy and efficiency of deep learning pipelines for automated contouring of organs in the mouse thorax.

Materials and Methods

We trained the 2D and 3D U-Net architectures of no-new-Net (nnU-Net) and AIMOS (i.e., current best performing algorithm for mouse segmentation) deep learning pipelines on 105 native micro-CT scans of mice, and we tested the trained models against an independent dataset (n=35, native CTs not included in training). Additionally, we also evaluated the segmentation performance on an external dataset (n=35, contrast-enhanced CTs), which do not share the same properties such as the mouse strain and image acquisition parameters as the training data. The quality of the automated contours was evaluated in terms of the mean surface distance (MSD) and 95% Hausdorff distance (95% HD). We also report the average preprocessing and inference times and the total runtime of each model.

Results

For the native CT dataset, all models of nnU-Net (3d_fullres, 3d_lowres, 3d_cascade, 2d) and AIMOS generated accurate contours of the heart, spinal cord, left and right lungs as shown in figure 1(a). They achieved an average MSD less than the in-plane voxel size of 0.14 mm while the average 95% HD were below 0.60 mm for all target organs except for the right lung segmentation of nnU-Net 2d. For the contrast-enhanced CTs, we chose to compare only the best performing 3D model of nnU-Net (3d_fullres) to the 2D models (nnU-Net 2d and AIMOS). Consistent for all organs, the nnU-Net 3d_fullres model showed superior segmentation performance (figure 2(a)). The 2D models generated incomplete contours and exhibited unacceptably large Hausdorff distances (> 1 mm). Although the 2D models are generally faster, all models took < 1 minute to generate contours, which is a huge improvement from the manual contouring time of ~40 minutes.

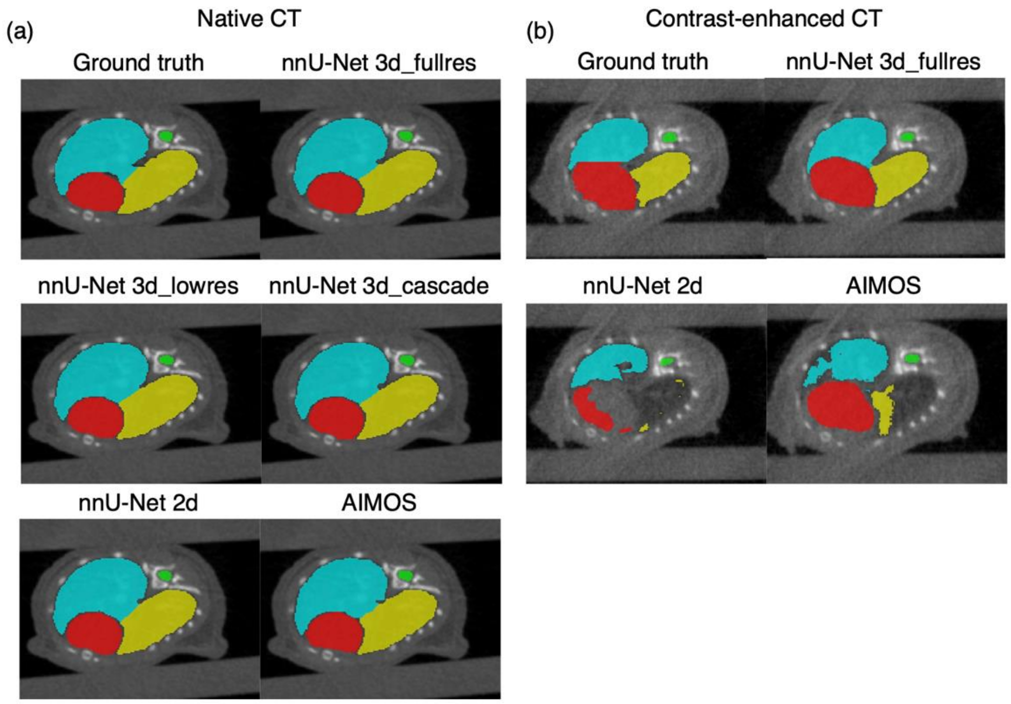

Figure 1. Example of automated contours for (a) native CTs and (b) contrast-enhanced CTs.

Table 1. Average preprocessing and inference times, and total runtimes in seconds.

Algorithm	Model	Preprocessing	Inference	Runtime
nnU-Net	3d_fullres	5	27	40
	3d_lowres	14	13	50
	3d_cascade (2^{nd} stage only)	17	28	52
	2d	5	7	21
AIMOS	UNet-768	6	2	20

Conclusion

We have shown that the nnU-Net 3d_fullres model outperforms the state-of-the-art AIMOS deep learning pipeline for mouse thoracic segmentation, and it offers a 98% reduction in contouring time compared to manual contouring. Our findings demonstrate the potential of integrating nnU-Net in routine practice to expedite irradiation and reduce the workload in preclinical facilities.