

# Supplementary material

## Swin SMT: Global Sequential Modeling for Enhancing 3D Medical Image Segmentation

Szymon Płotka<sup>1,2</sup>, Maciej Chrabaszcz<sup>2</sup>, and Przemyslaw Biecek<sup>2</sup>

<sup>1</sup> Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands

<sup>2</sup> Warsaw University of Technology, Warsaw, Poland

maciej.chrabaszcz.dokt@pw.edu.pl

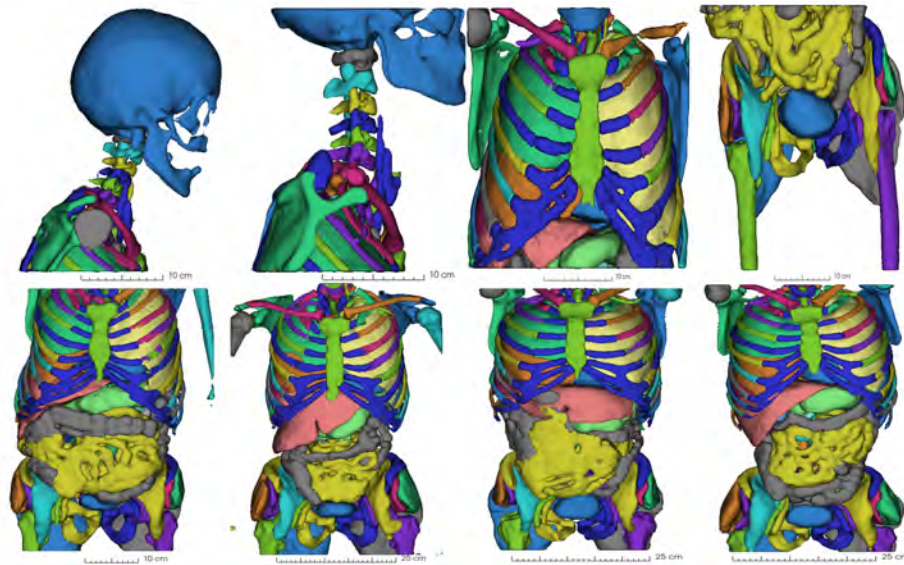


Fig. 1: We show partial and entire-body part CT predictions of Swin SMT. On the top, we show partial predictions of various body parts, including head-neck, chest, and pelvis parts, which can provide less robust predictions due to less contextual information. On the bottom, we show entire-body predictions, for which model is more robust due to the global and contextual sequential information provided by Swin SMT. We achieved significantly higher segmentation performance on the entire-body scans rather than partial. We demonstrate our predictions with real word scale in centimeters [cm].

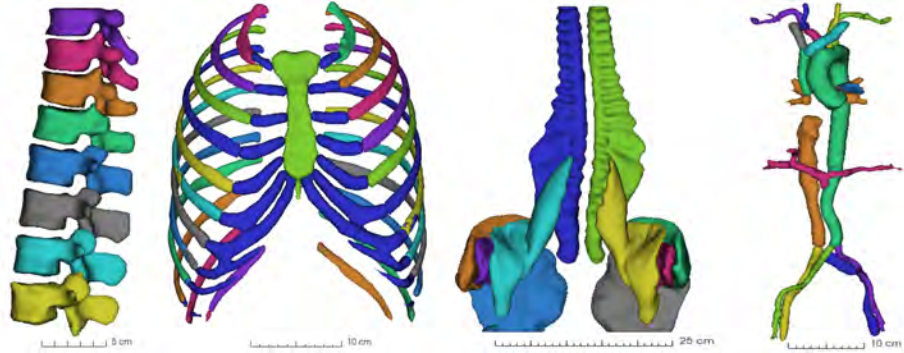


Fig. 2: We show top-performing predictions of Swin SMT for various subparts of the entire-body predictions. We provide the Dice Score Coefficient (in %) to show the segmentation performance. From the left: vertebrae (93.02%), ribs with sternum and costal cartilages (95.07%), muscles (92.34%), and vessels (86.00%). We demonstrate our predictions with real word scale in centimeters [cm].

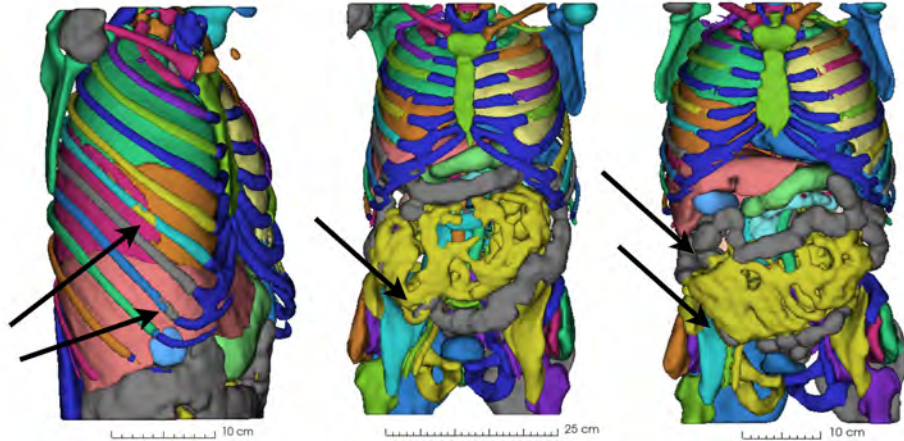


Fig. 3: We show the errors of the predictions of Swin SMT. While the quantitative segmentation performance is high (above 90% for each scan), we deal with some qualitative errors, as shown with black arrows. We demonstrate our predictions with real word scale in centimeters [cm].