



Supplementary Fig. 3. Comparison of the semantic segmentation model (U-Net) and the instance segmentation model (Mask R-CNN). (A) Original image of the lateral spine radiograph. (B) Ground truth (GT) segmentation by an experienced radiologist using 6-point labeling. (C) Semantic segmentation by U-Net, which failed to segment overlapped T11 and T12 vertebral bodies (VBs), independently. The segmentation was too fit to osteophytes and deformed VBs. (D) Instance segmentation by Mask R-CNN, with bounding boxes and contours in various colors, closely aligned with the GT and outperforming U-Net.