

# MedROV: Towards Real-Time Open-Vocabulary Detection Across Diverse Medical Imaging Modalities - Supplementary Material

**Analysis of Failure Cases:** Figure 1 illustrates failure modes observed in MedROV's predictions. Panel (a) shows a missed detection, where the model fails to identify visible organs such as the lungs, liver, vertebrae and ribs. Panels (b) and (c) depict duplicate detections, with multiple overlapping boxes assigned to the same left/right kidney. Panels (d) and (e) highlight left-right kidney confusion, where the model swaps the anatomical labels of the kidneys. These errors may arise from anatomical symmetry or ambiguous visual cues during training. Addressing such cases will be our future work.

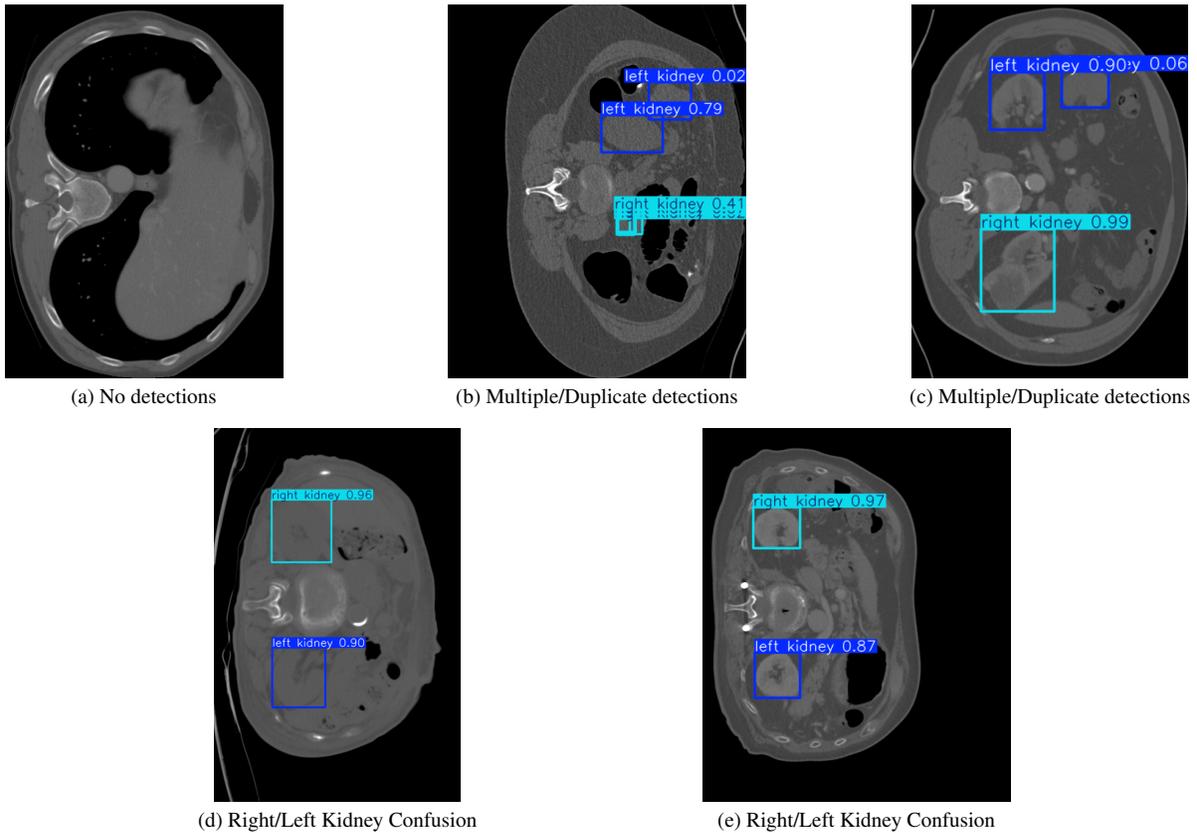


Figure 1. **Failure cases analysis.** Typical errors include: missed detections (a), duplicate boxes (b and c), left/right kidney confusion (d and e).