

**Supplementary Material: Learning to Segment Multiple Organs
from Multimodal Partially Labeled Datasets**

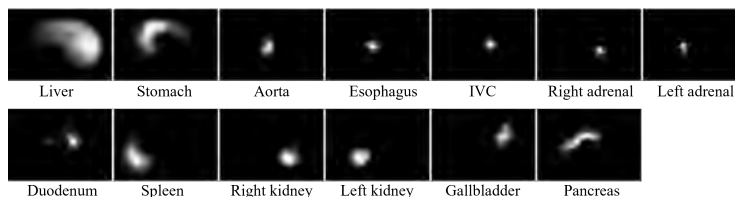


Fig. S1. Example a priori probabilistic atlases for 13 abdominal organs. IVC: inferior vena cava.

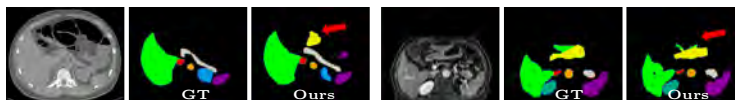


Fig. S2. Example segmentation failures (right: CT; left: MRI) with unsatisfactory regions marked by red arrows, likely due to mistaking for other organs and the complexity of abdominal structures.

Table S1. Sensitivity analysis of the performance concerning the variation of the a priori probabilistic atlas in DSC (mean \pm std%), where the atlas is constructed with 50% and all available training data for each organ. The performance is stable and no statistical significance is found for the comparison, suggesting that our framework is robust to the variation of the atlas.

Atlas construction data	Multimodal	
	Intramodal	Cross-modal
100%	83.4 \pm 16.9	74.6 \pm 18.2
50%	83.1 \pm 17.2	73.8 \pm 18.4

Table S2. Training-set pseudo-label quality comparison with two pseudo-label-based methods PaNN and U²PL in DSC (mean \pm std%). *: $p < 0.05$ for pairwise comparison with our method. The pseudo labels produced by our method are significantly better than those by PaNN and U²PL, especially in cross-modal circumstances.

	PaNN [40]	U ² PL [31]	Ours
Intramodal	83.2 \pm 14.1*	83.8 \pm 13.4*	85.2 \pm 15.1
Cross-modal	62.8 \pm 30.3*	62.6 \pm 29.3*	76.5 \pm 23.8

