Supplementary Appendix

Appendix

In the following section, we report the mean and standard deviation (std) of the Dice or clDice scores for the 15 tasks (Table A.1 and Table A.2), and the few-shot segmentation results of renal artery in "(mean)±(std)" form (Table A.3). We also show visualization examples of 3 organ segmentation tasks (Fig. A.1).

Table A.1. Quantitative comparison of Dice scores (%) for segmentation of 7 non-tubular-structure organs, and 1 tumor, shown as "(mean)±(std)". The sole-path model represents our model without the automatic pathway module trained in each task. We also show the number of training and testing cases ("training/testing") after the task name.

Method	Stomach (1,626/183)	Lung (345/20)	Ear (45/5)	Bladder (501/57)	Sacrum (266/35)	Cardiac (330/20)	Thyroid (334/38)	Lung tumor (57/6)
nnU-Net	90.91±10.80	**98.80±0.29**	86.05±2.43	90.79±7.80	92.57±16.17	87.86±10.19	**89.23±4.51**	77.12±11.08
Cascaded Vb-Net	91.59±14.43	98.14±0.35	80.99±5.84	86.13±17.18	90.54±16.07	91.26±5.49	83.00±7.35	59.11±27.86
Sole-path model	90.45±11.94	98.40±0.36	83.93±2.34	83.13±13.41	91.16±15.77	89.62±8.71	85.15±10.34	76.40±7.65
Ours	**92.59±8.29**	98.61±0.29	**86.29±1.79**	**91.68±5.62**	**94.86±1.32**	**93.01±3.48**	88.90±3.94	**80.75±8.71**

Table A.2. Quantitative comparison of Dice scores (%) and clDice scores (%) for segmentation of 7 tubular-structure organs, shown as "(mean)±(std)". We also show the number of training and testing cases ("training/testing") after the task name.

Metric	Method	Airway (108/20)	Coronary (6,065/155)	Pulmonary artery (1,153/20)	Pulmonary vein (352/20)	Aorta (1,694/20)	Lower limb artery (634/73)	Head and neck artery (3,664/729)
Dice (%)	nnU-Net	87.27±3.11	75.85±6.50	88.17±2.76	87.02±3.57	**95.49±1.15**	76.74±4.00	87.80±8.60
	Cascaded Vb-Net	85.70±2.16	**83.02±6.93**	87.50±2.61	85.70±4.44	89.54±2.84	79.73±5.12	93.19±4.11
	Sole-path model	86.79±3.58	79.55±5.02	87.71±2.75	84.07±7.22	90.90±3.15	78.86±4.62	93.50±4.20
	Ours	**87.60±2.90**	82.72±4.59	**88.83±2.80**	**87.16±4.42**	95.31±1.24	**80.13±4.42**	**94.45±2.74**
clDice (%)	nnU-Net	82.96±3.50	84.43±6.86	87.32±2.97	90.29±4.41	99.83±0.50	79.23±6.71	81.64±7.52
	Cascaded Vb-Net	78.49±7.01	88.10±7.12	85.88±3.30	87.59±5.86	96.46±2.58	75.80±7.87	91.29±3.42
	Sole-path model	83.10±5.66	86.68±5.89	85.86±4.13	88.38±4.34	94.92±6.67	78.37±7.06	91.70±3.84
	Ours	**84.47±3.37**	**90.72±5.10**	**88.58±2.46**	**92.08±2.71**	**99.88±0.35**	**82.38±6.21**	**93.28±2.74**

Table A.3. Comparison with other methods for the few-shot segmentation of renal artery, shown as "(mean)±(std)".

Training Setting	Method	Dice (%)	clDice (%)	Tsens (%)	Tprec (%)
1-shot	nnU-Net	82.56±11.52	56.17±15.02	41.88±15.18	**97.45±2.18**
	Cascaded Vb-Net	71.53±18.68	45.59±12.85	27.99±12.05	87.12±10.51
	Ours	**84.84±5.44**	**67.04±9.30**	**55.07±8.39**	87.30±14.47
5-shot	nnU-Net	**89.81±3.30**	77.40±8.58	65.45±11.24	**98.47±1.56**
	Cascaded Vb-Net	85.79±3.03	61.57±9.95	47.13±11.12	96.11±4.12
	Ours	88.64±3.23	**80.77±6.81**	**71.43±8.64**	94.30±6.92

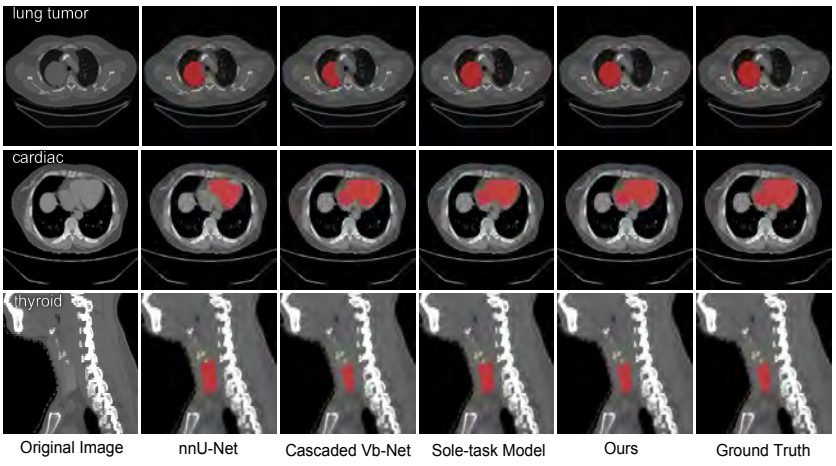

Fig. A.1. Examples of 3 organ segmentation tasks. The segmentation results or ground truth are shown in red masks.