Appendix

1 Qualitative results of models on the class-level continual multi-organ segmentation task

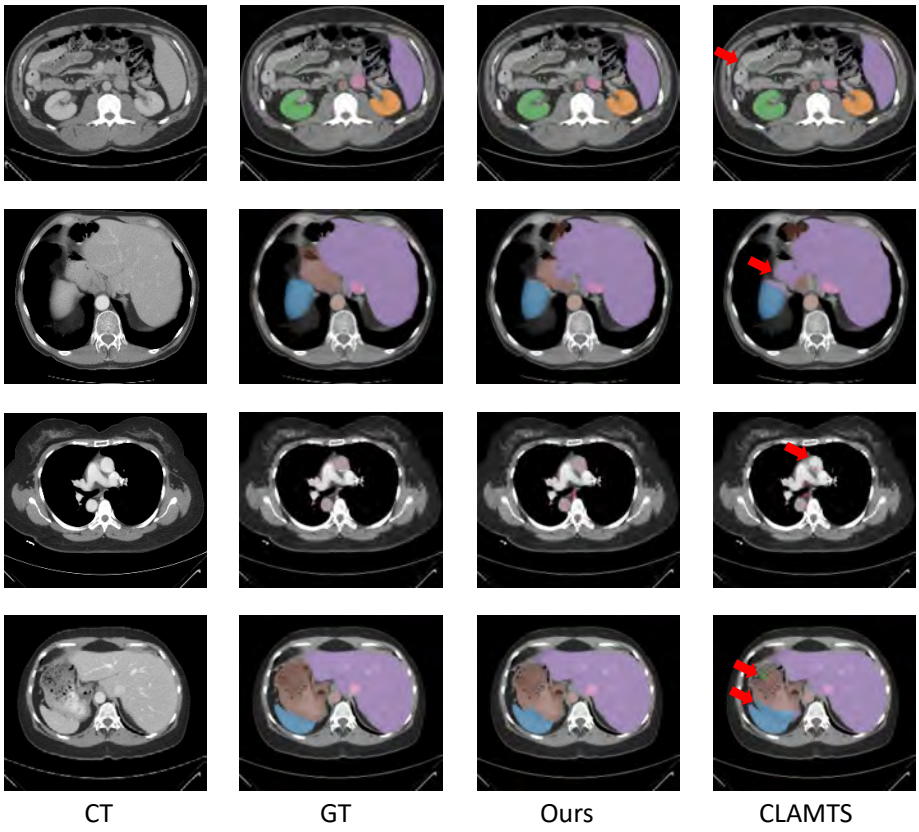

Fig. 1. The visualization comparison between our model and CLAMTS in continual learning step 2 on the BTCV dataset.

2 Quantitative results of models on the task-level continual medical image segmentation task

Table 1. Benchmark task-level continual learning methods. ∇, △ and □ represents ACDC, ISIC and COVID-19 CT dataset respectively. Red indicates the performance of the data trained in the final stage.

Model	#Param	low-rank	MoE	ACDC ∇	ISIC△	COVID-19 CT□
SOTAs	-	-	-	0.9146	0.8903	0.6820
∇ → △	88.1M	×	×	0.5019	0.9069	-
∇ → □	88.1M	×	×	0.3241	-	0.7244
△ → ∇	88.1M	×	×	0.9203	0.1347	-
△ → □	88.1M	×	×	-	0.5333	0.7214
□ → ∇	88.1M	×	×	0.9167	-	0.0003
□ → △	88.1M	×	×	-	0.9048	0.3818
∇ → △ → □	88.1M	×	×	0.0303	0.6360	0.7230
∇ → □ → △	88.1M	×	×	0.3908	0.9046	0.5020
△ → ∇ → □	88.1M	×	×	0.3746	0.0000	0.7223
△ → □ → ∇	88.1M	×	×	0.9192	0.4691	0.0000
□ → ∇ → △	88.1M	×	×	0.7094	0.8998	0.0417
□ → △ → ∇	88.1M	×	×	0.9201	0.1645	0.0000
∇ → △	3.4M	✓	✓	0.9208	0.9077	-
∇ → □	3.4M	✓	✓	0.9208	-	0.7368
△ → ∇	3.4M	✓	✓	0.9198	0.9063	-
△ → □	3.4M	✓	✓	-	0.9063	0.7434
□ → ∇	3.4M	✓	✓	0.9215	-	0.7327
□ → △	3.4M	✓	✓	-	0.9074	0.7327
∇ → △ → □	3.4M	✓	✓	0.9208	0.9077	0.7417
∇ → □ → △	3.4M	✓	✓	0.9208	0.9075	0.7368
△ → ∇ → □	3.4M	✓	✓	0.9198	0.9063	0.7428
△ → □ → ∇	3.4M	✓	✓	0.9210	0.9063	0.7434
□ → ∇ → △	3.4M	✓	✓	0.9215	0.9076	0.7327
□ → △ → ∇	3.4M	✓	✓	0.9212	0.9074	0.7327