

Table S 1 Segmentation performance comparison across different models (Scan Level). All models were trained using a 5-fold cross-validation schema, and scan-level evaluation metrics are reported on the test set as the mean of the fold-wise means $\pm$ the mean of the fold-wise standard deviation.

Model	Dice(%)	IoU(%)	95%HD	VS(%)
UNet	41.64 $\pm$ 31.63	31.39 $\pm$ 25.43	124.41 $\pm$ 50.90	61.30 $\pm$ 38.23
MultiView	40.95 $\pm$ 30.54	30.41 $\pm$ 24.41	121.76 $\pm$ 53.63	57.69 $\pm$ 38.81
nnUNet	68.82 $\pm$ 15.94	54.54 $\pm$ 16.92	110.27 $\pm$ 67.62	88.28 $\pm$ 11.42
SAM	43.63 $\pm$ 17.12	29.57 $\pm$ 13.80	136.69 $\pm$ 49.09	74.19 $\pm$ 20.59
SAM-Adapter	75.51$\pm$11.45	61.92$\pm$13.65	94.96$\pm$47.76	92.05$\pm$7.49

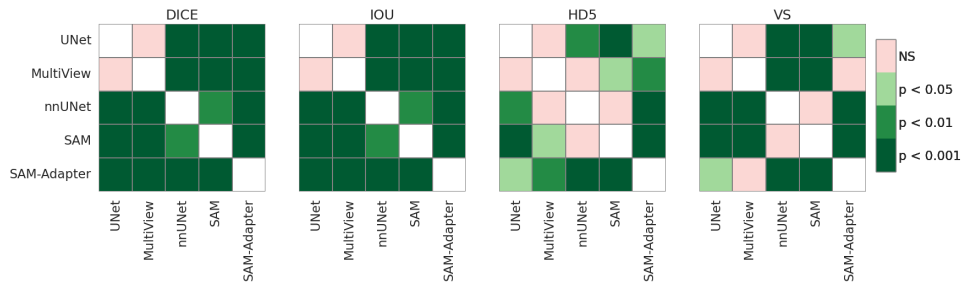


Figure S 1 Nemenyi post-hoc test assessing the statistical significance of model performance differences.

Table S 2 Optimization Experiments Based on SAM-Adpter Model (Scan Level). All the models were trained based on the 5-fold cross-validation schema, and the scan-level evaluation metrics are reported on the test set as the mean of the fold-wise means $\pm$ the mean of the fold-wise standard deviation.)

Model	Dice(%)	IoU(%)	95%HD	VS(%)
BCE	73.89 $\pm$ 10.35	59.59 $\pm$ 12.16	96.01 $\pm$ 49.43	91.25 $\pm$ 8.09
BCE(75%) + BL(25%)	73.81 $\pm$ 11.28	59.69 $\pm$ 13.22	90.84 $\pm$ 47.91	91.33 $\pm$ 7.78
BCE(50%) + BL(50%)	75.51 $\pm$ 11.45	61.92 $\pm$ 13.65	94.96 $\pm$ 47.76	92.05 $\pm$ 7.49
BCE(25%) + BL(75%)	74.22 $\pm$ 11.54	60.28 $\pm$ 13.63	94.53 $\pm$ 46.99	92.31 $\pm$ 6.83
Dice	74.47 $\pm$ 11.17	60.50 $\pm$ 13.14	85.23 $\pm$ 47.95	92.15 $\pm$ 7.70
Dice(50%) + BL(50%)	74.24 $\pm$ 11.57	60.30 $\pm$ 13.65	88.05 $\pm$ 49.65	92.75 $\pm$ 6.60
Focal Tversky(α =0.1)	69.68 $\pm$ 11.54	54.64 $\pm$ 12.66	105.57 $\pm$ 44.01	80.62 $\pm$ 12.68
Focal Tversky(α =0.2)	74.13 $\pm$ 10.56	59.94 $\pm$ 12.37	97.89 $\pm$ 51.94	86.24 $\pm$ 9.44
Focal Tversky(α =0.4)	75.65 $\pm$ 10.75	61.96 $\pm$ 13.03	92.15 $\pm$ 45.33	91.71 $\pm$ 6.43
Focal Tversky(α =0.6)	74.46 $\pm$ 11.96	60.64 $\pm$ 13.80	92.54 $\pm$ 47.93	91.51 $\pm$ 8.79
Adapter(M=20)	72.86 $\pm$ 13.16	58.86 $\pm$ 14.79	93.65 $\pm$ 46.02	91.01 $\pm$ 9.73
Adapter(M=50)	74.68 $\pm$ 11.37	60.82 $\pm$ 13.49	92.18 $\pm$ 52.56	92.67 $\pm$ 6.27
Adapter(M=100)	75.51 $\pm$ 11.45	61.92 $\pm$ 13.65	94.96 $\pm$ 47.76	92.05 $\pm$ 7.49
Adapter (M=200)	74.71 $\pm$ 11.59	60.90 $\pm$ 13.66	89.29 $\pm$ 47.28	92.38 $\pm$ 6.64
No CA*, No Pretrain	71.32 $\pm$ 12.52	56.84 $\pm$ 14.23	106.99 $\pm$ 50.89	91.11 $\pm$ 6.23
No CA, Pretrain w/ BHSD	72.15 $\pm$ 12.59	57.83 $\pm$ 14.11	101.31 $\pm$ 48.34	91.27 $\pm$ 6.81
No CA, Pretrain w/ Instance	69.96 $\pm$ 13.13	55.30 $\pm$ 14.51	109.14 $\pm$ 49.43	90.29 $\pm$ 6.52
No CA, Pretrain w/ PhysioNet	71.18 $\pm$ 12.05	56.53 $\pm$ 13.58	110.97 $\pm$ 52.89	91.91 $\pm$ 5.47
No CA, Pretrain w/ All	73.35 $\pm$ 11.74	59.19 $\pm$ 13.45	97.53 $\pm$ 52.78	91.65 $\pm$ 6.12
w/ CA, No pre-train	75.51 $\pm$ 11.45	61.92 $\pm$ 13.65	94.96 $\pm$ 47.76	92.05 $\pm$ 7.49

* CA: Contrast Adjustment

* Focal Tversky: Other parameters were kept constant with γ =1 and ϵ =0.01

Table S 3 Friedman’s Test on different model architectures(Scan Level)

Metrics Model	Dice		IoU		95%HD		VS(%)	
	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
UNet								
MultiView	47.21	3.13e-10***	47.21	3.13e-10***	14.92	1.89e-3***	31.16	7.87e-7***
nnUNet								
SAM								
SAM-Adapter								

* p-value <0.05
** p-value <0.01
*** p-value <0.001

Table S 4 Friedman's Test on Model Optimization Experiments (Scan Level)

Metrics Model	Dice		IoU		95%HD		VS(%)	
	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
BCE								
BCE(75%) + BL(25%)								
BCE(50%) + BL(50%)								
BCE(25%) + BL(75%)								
Dice	5.47	0.14	5.47	0.14	1.26	0.73	0.57	0.90
Dice(50%) + BL(50%)								
Focal Tversky($\alpha=0.1$)								
Focal Tversky($\alpha=0.2$)								
Focal Tversky($\alpha=0.4$)								
Focal Tversky($\alpha=0.6$)								
Adapter(M=20)								
Adapter(M=50)								
Adapter(M=100)	4.73	0.14	4.73	0.19	0.14	0.99	0.32	0.96
Adapter (M=200)								
No CA*, No Pretrain								
No CA, Pretrain w/ BHSD								
No CA, Pretrain w/ Instance	5.40	0.14	5.40	0.14	1.29	0.73	1.44	0.70
No CA, Pretrain w/ BC1HM								
No CA, Pretrain w/ All								
w/ CA, No pre-train								

* CA: Contrast Adjustment

* Focal Tversky: Other parameters were kept constant with $\gamma=1$ and $\epsilon=0.01$