

Table S1. Segmentation results on AbdomenMRI dataset with standard deviation. The results of nnU-Net, SegResNet, UNETR, SwinUNETR, and U-Mamba were referenced from [18]. *: Deep supervision was disabled and we extend the training epochs to 200.

Methods	param	FLOPs	DSC	NSD
<i>CNN-based</i>				
nnU-Net	33M	23.3G	0.7450 ± 0.1117	0.8153 ± 0.1145
SegResNet	6M	24.5G	0.7317 ± 0.1379	0.8034 ± 0.1386
<i>Transformer-based</i>				
UNETR	87M	42.1G	0.5747 ± 0.1672	0.6309 ± 0.1858
SwinUNETR	25M	27.9G	0.7028 ± 0.1348	0.7669 ± 0.1442
nnFormer	60M	50.2G	0.7297 ± 0.1486	0.7963 ± 0.1322
<i>Mamba-based</i>				
U-Mamba_Bot	63M	45.7G	0.7588 ± 0.1051	0.8285 ± 0.1074
U-Mamba_Enc	67M	49.9G	0.7625 ± 0.1082	0.8327 ± 0.1087
<i>w/o ImageNet-based pretraining</i>				
Swin-UMamba	60M	68.0G	0.7054 ± 0.1387	0.7647 ± 0.1455
Swin-UMamba†	28M	18.9G	$0.6653 \pm 0.1123^*$	$0.7312 \pm 0.1199^*$
<i>w/ ImageNet-based pretraining</i>				
Swin-UMamba	60M	68.0G	0.7760 ± 0.0955	0.8421 ± 0.0964
Swin-UMamba†	28M	18.9G	0.7705 ± 0.0963	0.8376 ± 0.0981

Table S2. Segmentation results on Endoscopy and Microscopy dataset with standard deviation. The results of nnU-Net, SegResNet, UNETR, SwinUNETR, and U-Mamba were referenced from [18].

Methods	Endoscopy		Microscopy
	DSC	NSD	F1
<i>CNN-based</i>			
nnU-Net	0.6264 ± 0.3024	0.6421 ± 0.3074	0.5383 ± 0.2657
SegResNet	0.5820 ± 0.3268	0.5968 ± 0.3303	0.5411 ± 0.2633
<i>Transformer-based</i>			
UNETR	0.5017 ± 0.3201	0.5168 ± 0.3235	0.4357 ± 0.2572
SwinUNETR	0.5528 ± 0.3089	0.5683 ± 0.3123	0.3967 ± 0.2621
nnFormer	0.6135 ± 0.2763	0.6228 ± 0.2832	0.5332 ± 0.2543
<i>Mamba-based</i>			
U-Mamba_Bot	0.6540 ± 0.3008	0.6692 ± 0.3050	0.5389 ± 0.2817
U-Mamba_Enc	0.6303 ± 0.3067	0.6451 ± 0.3104	0.5607 ± 0.2784
<i>w/o ImageNet-based pretraining</i>			
Swin-UMamba	0.5483 ± 0.3047	0.5632 ± 0.3085	0.4561 ± 0.2806
Swin-UMamba†	0.6402 ± 0.3260	0.6547 ± 0.3301	0.5186 ± 0.2727
<i>w/ ImageNet-based pretraining</i>			
Swin-UMamba	0.6767 ± 0.2866	0.6922 ± 0.2918	0.5806 ± 0.2297
Swin-UMamba†	0.6783 ± 0.2969	0.6933 ± 0.3011	0.5982 ± 0.2364

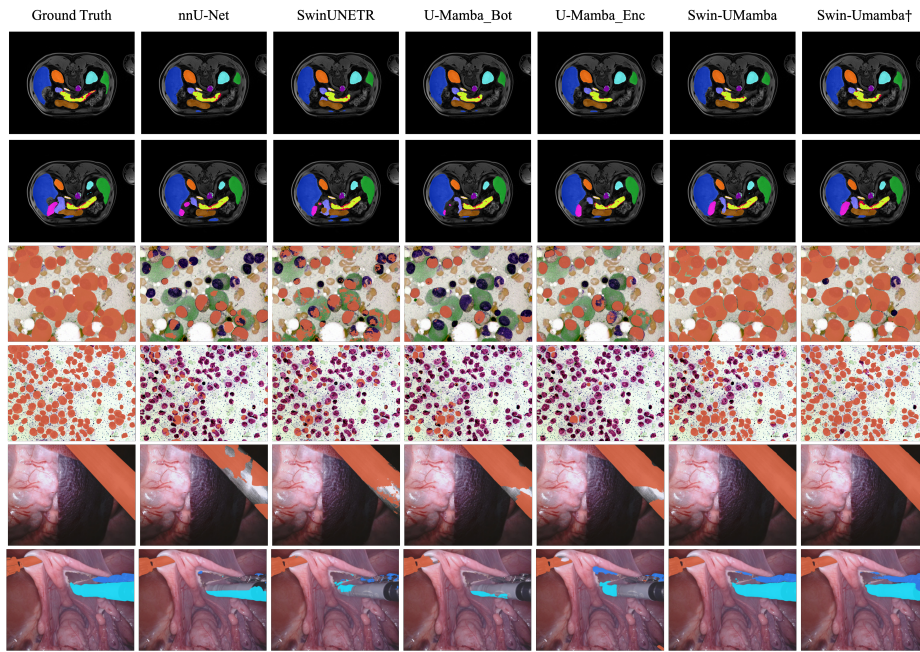


Fig. S1. Result visualization on AbdomenMRI (top two rows), Endoscopy (middle two rows), and Microscopy (bottom two rows) dataset.

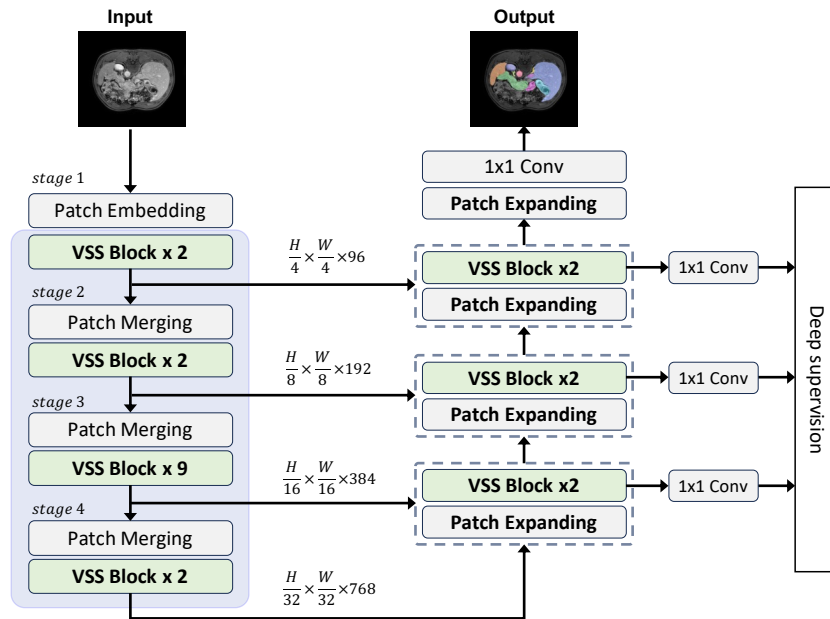


Fig. S2. The overall architecture of Swin-UMamba†.