Appendix

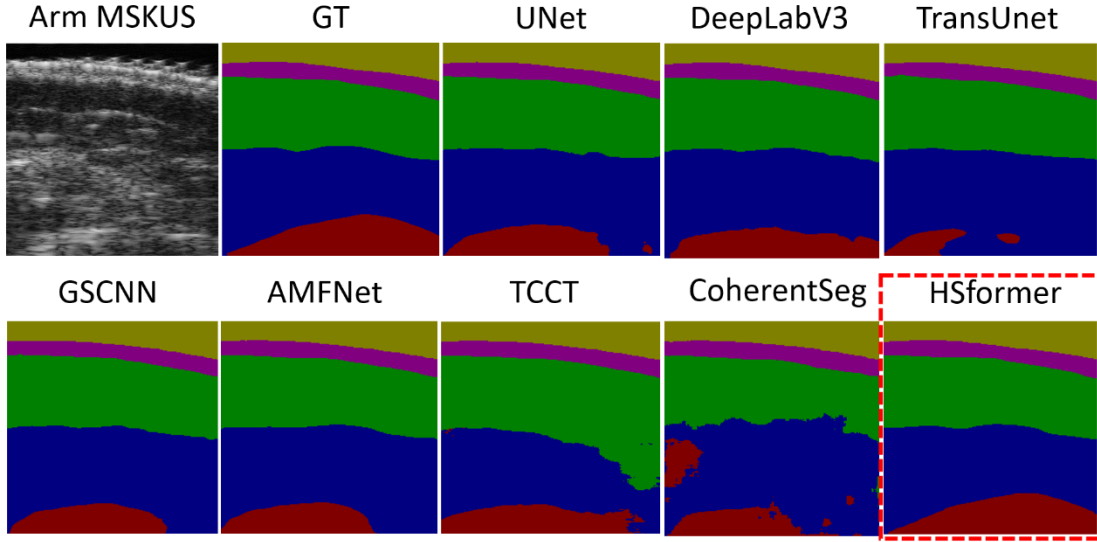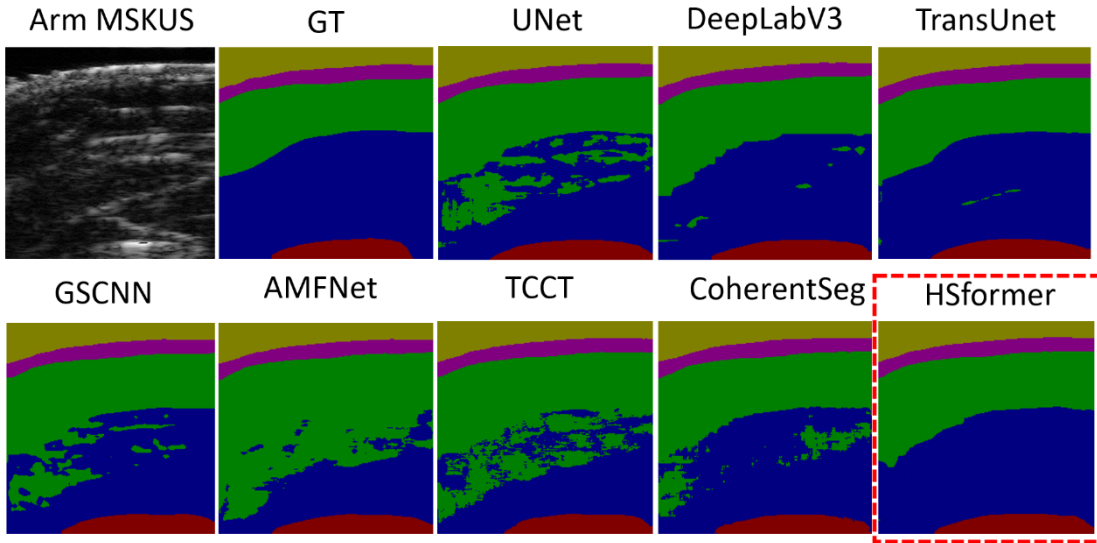**A1.** Visualization of predicted segmentation maps obtained from all models

(a)

(b)

Figure A1. Two cases are selected to demonstrate the segmentation performance of seven comparison models and the proposed HSformer.

A2. Qualitative and quantitative results compared to the SAM (Segment Anything) model

SAM-MP: Prompt includes a randomly marked background point as a negative point, and multiple foreground points marked as positive points.

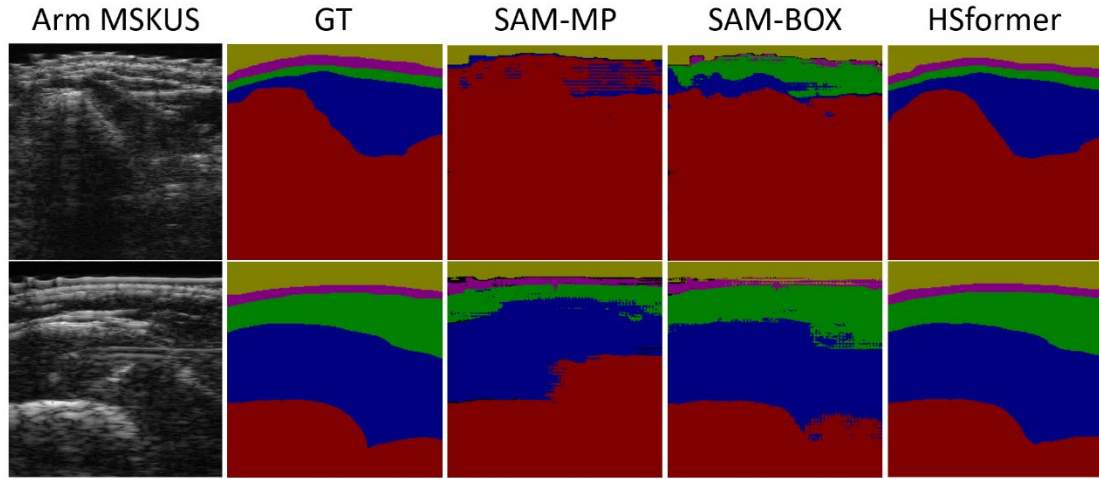SAM-BOX: The bounding box of the GT mask was directly treated as a prompt.

Figure A2. The segmentation performance of SAM-MP, SAM-BOX and the proposed HSformer.

Table A1. Segmentation results of SAM-MP, SAM-BOX and the proposed HSformer.

layer	gel				skin				subcutaneous fat				muscles				bones				mean			
structures	DSC	JI	HD	ASD	DSC	JI	HD	ASD	DSC	JI	HD	ASD	DSC	JI	HD	ASD	DSC	JI	HD	ASD	DSC	JI	HD	ASD
SAM-MP	0.70	0.57	28.72	8.33	0.41	0.29	50.04	12.58	0.62	0.48	58.13	16.35	0.71	0.58	70.8	17.92	0.70	0.57	67.23	17.83	0.52	0.41	45.82	12.16
SAM-BOX	0.85	0.75	11.76	3.13	0.62	0.48	14.69	5.09	0.79	0.69	25.08	6.76	0.79	0.67	56.52	11.95	0.87	0.79	24.99	5.12	0.78	0.68	26.61	6.41
HSformer	**0.98**	**0.95**	**2.27**	**0.48**	**0.87**	**0.78**	**2.77**	**1.06**	**0.90**	**0.83**	**13.21**	**3.19**	**0.89**	**0.81**	**25.02**	**5.16**	**0.89**	**0.82**	**19.76**	**3.90**	**0.91**	**0.84**	**12.61**	**2.76**

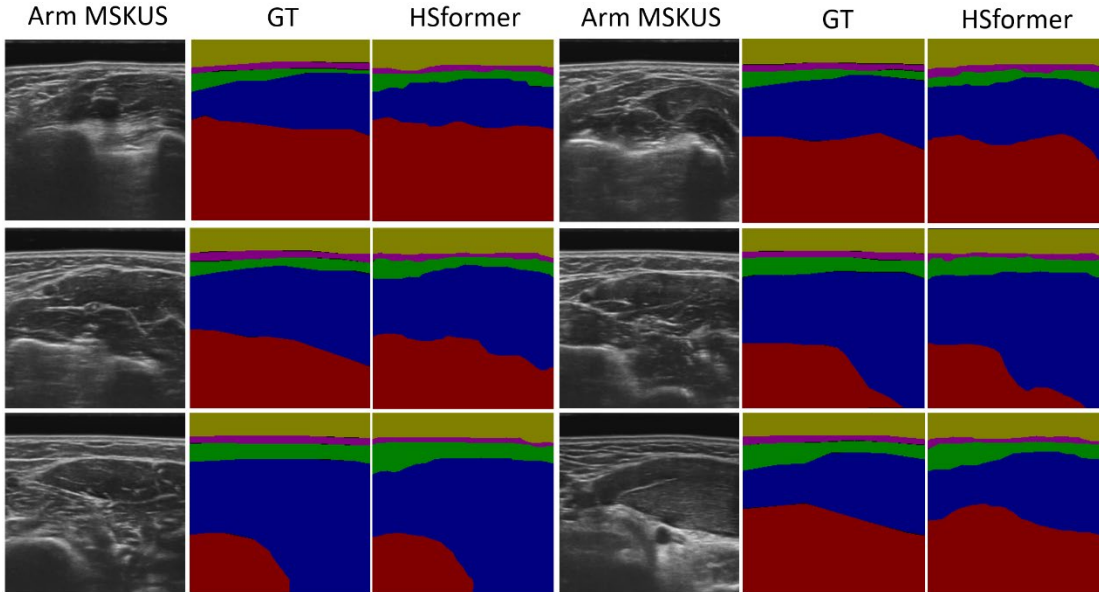**A3**. In-house Small-scale Arm MSKUS Test

Figure A3. Cases acquired by directly using HSformer on untrained datasets.