This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

UltraPIPS: Supplementary Material

Tal Grutman and Tali Ilovitsh

School of Biomedical Engineering, Tel Aviv University, Tel Aviv, Israel
ilovitsh@tauex.tau.ac.il

Table S1. Spearman correlation of EchoPrime view classifier confidence with distance metrics on augmented data. For each augmentation type, the augmentation is applied with increased strength. Correlations are measured between view classifier log-softmax output at the correct view class and each of the candidate distance metrics, averaged across frames. In each column, the highest correlation is displayed in **bold**.

	CAMUS				EchoNet				Overall
Model	Depth	Rotation	Sector Width	Translation	Depth	Rotation	Sector Width	Translation	
L_2	0.57 ± 0.25	0.36 ± 0.29	0.58 ± 0.27	0.26 ± 0.28	0.50 ± 0.27	0.40 ± 0.29	0.53 ± 0.27	0.37 ± 0.25	0.45 ± 0.28
SSIM	0.42 ± 0.26	0.20 ± 0.28	0.51 ± 0.28	0.18 ± 0.25	0.24 ± 0.27	0.09 ± 0.28	0.23 ± 0.27	0.02 ± 0.22	0.20 ± 0.27
AlexNet	0.58 ± 0.25	0.45 ± 0.26	**0.62 ± 0.25**	0.27 ± 0.27	0.54 ± 0.28	0.51 ± 0.29	0.56 ± 0.27	0.39 ± 0.25	0.49 ± 0.28
VGG-16	0.54 ± 0.25	0.37 ± 0.27	0.61 ± 0.26	0.25 ± 0.27	**0.55 ± 0.27**	0.52 ± 0.29	0.57 ± 0.27	0.39 ± 0.25	0.48 ± 0.28
ViT	0.52 ± 0.26	0.38 ± 0.26	0.62 ± 0.26	0.24 ± 0.27	0.53 ± 0.27	0.50 ± 0.28	0.55 ± 0.27	0.38 ± 0.25	0.47 ± 0.28
SwinViT	0.44 ± 0.26	0.34 ± 0.29	0.57 ± 0.27	0.13 ± 0.24	0.35 ± 0.25	0.39 ± 0.29	0.49 ± 0.27	0.27 ± 0.23	0.37 ± 0.28
CLIP	0.51 ± 0.26	0.39 ± 0.27	0.57 ± 0.27	0.23 ± 0.26	0.48 ± 0.27	0.41 ± 0.30	0.48 ± 0.26	0.35 ± 0.24	0.43 ± 0.28
RadImageNet	**0.61 ± 0.26**	**0.49 ± 0.26**	0.62 ± 0.26	0.28 ± 0.28	0.50 ± 0.26	0.48 ± 0.29	0.48 ± 0.27	0.35 ± 0.25	0.47 ± 0.29
MedSAM	0.58 ± 0.26	0.36 ± 0.30	0.58 ± 0.27	0.24 ± 0.28	0.46 ± 0.27	0.31 ± 0.29	0.49 ± 0.27	0.34 ± 0.25	0.42 ± 0.29
BiomedCLIP	0.59 ± 0.26	0.48 ± 0.26	0.61 ± 0.27	0.26 ± 0.28	0.51 ± 0.27	0.50 ± 0.29	0.55 ± 0.27	0.38 ± 0.25	0.48 ± 0.29
USFM	0.56 ± 0.25	0.37 ± 0.28	0.56 ± 0.26	0.27 ± 0.28	0.54 ± 0.27	0.42 ± 0.29	0.53 ± 0.27	0.37 ± 0.25	0.45 ± 0.28
TUSA	0.59 ± 0.26	0.48 ± 0.25	0.60 ± 0.27	0.27 ± 0.28	0.53 ± 0.27	**0.53 ± 0.29**	**0.57 ± 0.26**	**0.40 ± 0.25**	**0.50 ± 0.28**
Ultrasound-CLIP	0.58 ± 0.25	0.48 ± 0.25	0.62 ± 0.26	**0.28 ± 0.28**	0.50 ± 0.27	0.51 ± 0.29	0.55 ± 0.26	0.39 ± 0.25	0.49 ± 0.28

Table S2. Spearman correlation of UltraSam dice scores with distance metrics on augmented data. For each augmentation type, the augmentation is applied with increased strength. Correlations are measured between dice score and each of the candidate distance metrics, averaged across frames. In each column, the highest correlation is displayed in **bold**.

	CAMUS				EchoNet				Overall
Model	Depth	Rotation	Sector Width	Translation	Depth	Rotation	Sector Width	Translation	
L_2	0.04 ± 0.25	0.46 ± 0.29	0.45 ± 0.28	0.38 ± 0.25	0.10 ± 0.25	0.39 ± 0.30	0.31 ± 0.28	0.36 ± 0.25	0.30 ± 0.28
SSIM	0.01 ± 0.22	0.20 ± 0.27	0.39 ± 0.28	0.25 ± 0.24	0.06 ± 0.23	0.04 ± 0.30	0.14 ± 0.26	0.01 ± 0.22	0.12 ± 0.26
AlexNet	0.04 ± 0.25	0.67 ± 0.26	**0.51 ± 0.28**	**0.40 ± 0.25**	**0.13 ± 0.27**	**0.57 ± 0.29**	**0.35 ± 0.29**	**0.39 ± 0.25**	**0.37 ± 0.29**
VGG-16	0.02 ± 0.23	0.50 ± 0.28	0.48 ± 0.28	0.35 ± 0.24	**0.13 ± 0.27**	0.55 ± 0.29	**0.35 ± 0.28**	0.38 ± 0.25	0.34 ± 0.28
ViT	0.02 ± 0.24	0.59 ± 0.27	0.48 ± 0.28	0.37 ± 0.25	0.12 ± 0.26	0.54 ± 0.30	0.34 ± 0.28	0.37 ± 0.25	0.35 ± 0.28
SwinViT	0.04 ± 0.24	0.52 ± 0.29	0.43 ± 0.27	0.27 ± 0.24	0.10 ± 0.24	0.42 ± 0.30	0.30 ± 0.27	0.25 ± 0.23	0.28 ± 0.28
CLIP	0.03 ± 0.24	0.58 ± 0.27	0.45 ± 0.29	0.33 ± 0.24	0.10 ± 0.26	0.45 ± 0.30	0.29 ± 0.27	0.31 ± 0.23	0.31 ± 0.28
RadImageNet	0.03 ± 0.26	**0.75 ± 0.24**	0.50 ± 0.29	**0.40 ± 0.25**	0.11 ± 0.25	0.53 ± 0.29	0.29 ± 0.28	0.34 ± 0.25	0.36 ± 0.29
MedSAM	0.03 ± 0.25	0.48 ± 0.29	0.46 ± 0.29	0.38 ± 0.25	0.09 ± 0.25	0.34 ± 0.29	0.28 ± 0.28	0.33 ± 0.25	0.29 ± 0.28
BiomedCLIP	0.04 ± 0.26	0.73 ± 0.24	0.47 ± 0.28	0.39 ± 0.25	0.10 ± 0.26	0.55 ± 0.30	0.33 ± 0.28	0.36 ± 0.24	0.36 ± 0.29
USFM	0.04 ± 0.25	0.53 ± 0.28	0.45 ± 0.28	0.37 ± 0.25	0.12 ± 0.26	0.46 ± 0.30	0.32 ± 0.28	0.37 ± 0.25	0.32 ± 0.28
TUSA	**0.05 ± 0.26**	0.73 ± 0.24	0.47 ± 0.28	**0.40 ± 0.26**	0.11 ± 0.26	0.56 ± 0.30	0.34 ± 0.28	**0.39 ± 0.25**	**0.37 ± 0.29**
Ultrasound-CLIP	0.04 ± 0.25	0.74 ± 0.23	0.48 ± 0.28	0.39 ± 0.25	0.10 ± 0.26	0.56 ± 0.30	0.33 ± 0.28	0.38 ± 0.24	**0.37 ± 0.29**

Table S3. Image reconstruction metrics of INRs trained with each individual distance metric. For each metric, the best performer is displayed in **bold**.

		↑ SSIM	↓ HFEN	↓ GLCM Contrast	↓ GLCM Dissimilarity	↓ GLCM Homogeneity	↓ GLCM Energy	↓ GLCM Correlation
USOVA	L_2	0.76 ± 0.09	0.34 ± 0.04	2.84 ± 1.92	0.45 ± 0.26	0.07 ± 0.04	0.02 ± 0.02	0.05 ± 0.04
	AlexNet	0.59 ± 0.12	0.69 ± 0.13	1.77 ± 1.39	0.22 ± 0.16	0.03 ± 0.02	**0.01 ± 0.01**	0.03 ± 0.02
	VGG-16	0.61 ± 0.12	0.61 ± 0.10	2.34 ± 1.25	0.30 ± 0.14	0.04 ± 0.02	**0.01 ± 0.01**	0.04 ± 0.02
	ViT	0.39 ± 0.12	1.26 ± 0.23	1.18 ± 1.10	0.15 ± 0.12	0.03 ± 0.03	0.05 ± 0.04	**0.01 ± 0.01**
	SwinViT	0.28 ± 0.10	1.59 ± 0.29	1.59 ± 1.37	0.29 ± 0.19	0.08 ± 0.05	0.09 ± 0.06	0.02 ± 0.02
	CLIP	0.36 ± 0.11	1.44 ± 0.25	2.18 ± 1.68	0.25 ± 0.20	0.04 ± 0.03	0.05 ± 0.04	0.03 ± 0.03
	RadImageNet	0.42 ± 0.12	1.25 ± 0.22	1.46 ± 1.26	0.18 ± 0.15	**0.02 ± 0.02**	0.02 ± 0.03	0.02 ± 0.02
	BiomedCLIP	0.52 ± 0.13	0.87 ± 0.16	2.13 ± 1.40	0.27 ± 0.16	0.03 ± 0.02	**0.01 ± 0.01**	0.04 ± 0.02
	USFM	**0.79 ± 0.09**	**0.30 ± 0.04**	1.46 ± 1.21	0.23 ± 0.15	0.04 ± 0.02	0.02 ± 0.02	0.03 ± 0.02
	TUSA	0.68 ± 0.11	0.50 ± 0.08	3.56 ± 2.10	0.52 ± 0.28	0.08 ± 0.04	**0.01 ± 0.01**	0.06 ± 0.04
	Ultrasound-CLIP	0.62 ± 0.13	0.58 ± 0.11	**0.87 ± 0.92**	**0.12 ± 0.11**	**0.02 ± 0.01**	**0.01 ± 0.01**	**0.01 ± 0.01**
MicroSegNet	L_2	0.65 ± 0.04	0.37 ± 0.04	5.42 ± 1.84	0.62 ± 0.21	0.06 ± 0.03	0.02 ± 0.01	0.09 ± 0.02
	AlexNet	0.45 ± 0.08	1.12 ± 0.20	3.69 ± 1.66	0.32 ± 0.16	0.02 ± 0.01	0.02 ± 0.01	0.06 ± 0.01
	VGG-16	0.50 ± 0.08	0.87 ± 0.16	4.18 ± 1.73	0.38 ± 0.17	0.02 ± 0.01	**0.00 ± 0.01**	0.06 ± 0.02
	ViT	0.32 ± 0.10	1.53 ± 0.25	2.36 ± 1.59	0.17 ± 0.13	0.04 ± 0.03	0.04 ± 0.02	0.03 ± 0.02
	SwinViT	0.25 ± 0.08	1.77 ± 0.25	4.52 ± 2.40	0.31 ± 0.21	0.05 ± 0.04	0.06 ± 0.03	0.05 ± 0.02
	CLIP	0.27 ± 0.09	1.76 ± 0.26	3.28 ± 2.12	0.24 ± 0.17	0.06 ± 0.04	0.06 ± 0.03	0.04 ± 0.02
	RadImageNet	0.38 ± 0.09	1.48 ± 0.24	4.27 ± 1.69	0.43 ± 0.17	0.03 ± 0.02	0.01 ± 0.01	0.07 ± 0.02
	BiomedCLIP	0.41 ± 0.09	1.25 ± 0.21	3.47 ± 1.88	0.30 ± 0.18	0.02 ± 0.01	0.01 ± 0.01	0.06 ± 0.02
	USFM	**0.66 ± 0.04**	**0.36 ± 0.05**	1.95 ± 0.95	0.18 ± 0.09	**0.01 ± 0.01**	0.01 ± 0.01	0.04 ± 0.01
	TUSA	0.52 ± 0.07	0.79 ± 0.13	6.64 ± 2.33	0.71 ± 0.26	0.07 ± 0.03	0.02 ± 0.01	0.11 ± 0.03
	Ultrasound-CLIP	0.47 ± 0.08	0.97 ± 0.21	**1.15 ± 0.78**	**0.09 ± 0.06**	0.02 ± 0.01	0.01 ± 0.01	**0.02 ± 0.01**