Supplementary material for Paper ID: 1343

(a) Mask Ratio

Mask Ratio	R@0.3
0.3	**0.61**
0.5	0.57
0.8	0.44

(b) Coefficient Decay

Strategy	R@0.3
No decay	0.65
Linear	0.72
Hard	**0.74**

(c) Training Data

Training Data	R@0.3
Source	0.63
Target	0.71
Source and Target	**0.74**

(d) Annealing

Annealing	R@0.3
Step	0.
Cyclic	0.71
Cosine	**0.74**

Table S1: The MAE ablation study shows that a low initial mask ratio, decaying λ_{mask}, and training on both source and target features improve performance, with D-MASTER reducing overfitting.

Hyper-parameter	Description	INBreast	DDSM
N_c	Number of categories for classification head	2	2
N_l^{enc}	Number of encoder layers	8	6
N_l^{dec}	Number of decoder layers	8	6
N_l^{aux}	Number of MAE auxiliary decoder layers	2	2
N_q^{dec}	Number of queries for decoder	300	300
N_q^{aux}	Number of queries for MAE auxiliary decoder	882	882
H	Number of hidden dimension for deformable attention	256	256
F	Number of feedforward dimension for deformable attention	1024	1024
L	Number of feature levels for deformable attention	4	4
M	Number of heads for deformable attention	8	8
K	Number of reference points for each attention head	4	4
B	Batch Size during training	16	16
lr	Learning rate for modules except backbone and projection	2×10^{-4}	2×10^{-4}
lr_{bac}	Learning rate for backbone and projection modules	2×10^{-5}	2×10^{-5}
β_{bac}	Coefficient of discrimination loss after backbone $L_{\text{bac}}^{\text{dis}}$	0.3	0.3
β_{enc}	Coefficient of discrimination loss after encoder $L_{\text{enc}}^{\text{dis}}$	1.0	1.0
β_{dec}	Coefficient of discrimination loss after decoder $L_{\text{enc}}^{\text{dis}}$	1.0	1.0
λ_{unsup}	Coefficient of unsupervised loss $\mathcal{L}_{unsup}$	1.0	1.0
λ_{mask}	Coefficient of supervised loss $\mathcal{L}_{sup}$	1.0	1.0
γ	EMA update ratio	0.9996	0.9996
μ_t	Initial Mask ratio in MAE branch	0.2	0.3
η	Initial step for annealing	0.2	0.3
η_{min}^i	Minimum jump in mask annealing step η	0.05	0.05
η_{max}^i	Maximum jump in mask annealing step η	0.15	0.15
C_s	Soft Confidence Metric	0.15	0.20
C_h	Hard Confidence Metric	0.80	0.90
E_{pre}	MAE branch with source data training epoch number	88	87
E_{teach}	Teacher-student training epoch number	84	76
E_{decay}	After Edecay epochs in teaching stage, we drop the MAE branch	30	10
E_{reinit}	Re-initialization epoch for selective retraining	40	20

Table S2: Below are the detailed hyper-parameters corresponding to each benchmark, with the source dataset as the In-house dataset

Datasets	Model Name	Venue	R@0.05	R@0.1	R@0.3	R@0.5	R@1.0	R@2.0	Accuracy	F1-score
In-house to DDSM [18]	SFA[36]	MM'21	0.01	0.01	0.05	0.07	0.11	0.313	0.216	0.329
	UMT[6]	CVPR'21	0.0	0.01	0.04	0.07	0.09	0.13	0.261	0.362
	D-Adapt[14]	ICLR'22	0.0	0.02	0.06	0.09	0.10	0.13	0.382	0.215
	AT[21]	CVPR'22	0.01	0.03	0.08	0.10	0.15	0.21	0.216	0.311
	H2FA[38]	CVPR'22	0.02	0.03	0.06	0.10	0.12	0.17	0.371	0.315
	AQT[13]	IJCAI'22	0.01	0.03	0.07	0.13	0.15	0.18	0.412	0.398
	HT[7]	CVPR'23	0.03	0.05	0.08	0.10	0.13	0.15	0.362	0.362
	ConfMIX[24]	WACV'23	0.02	0.04	0.09	0.12	0.16	0.19	0.336	0.412
	CLIPGAP[34]	CVPR'23	0.01	0.03	0.07	0.11	0.15	0.16	0.336	0.458
	MRT[43]	ICCV'23	0.03	0.04	0.09	0.12	0.17	0.21	0.421	0.587
	Ours	-	**0.02**	**0.05**	**0.12**	**0.17**	**0.29**	**0.49**	**0.561**	**0.613**
RSNA-BSD1K to In-house	SFA[36]	MM'21	0.03	0.07	0.13	0.18	0.27	0.31	0.629	0.210
	UMT[6]	CVPR'21	0.01	0.04	0.09	0.15	0.19	0.23	0.568	0.193
	D-Adapt[14]	ICLR'22	0.03	0.06	0.11	0.18	0.25	0.31	0.668	0.241
	AT[21]	CVPR'22	0.10	0.28	0.37	0.45	0.51	0.66	0.725	**0.319**
	H2FA[38]	CVPR'22	0.03	0.06	0.14	0.17	0.21	0.24	0.591	0.274
	AQT[13]	IJCAI'22	0.01	0.05	0.08	0.11	0.17	0.20	0.527	0.230
	HT[7]	CVPR'23	0.02	0.10	0.17	0.24	0.33	0.41	0.710	0.291
	ConfMIX[24]	WACV'23	0.03	0.09	.0.16	0.28	0.35	0.39	0.622	0.263
	CLIPGAP[34]	CVPR'23	0.04	0.08	0.23	0.36	0.64	0.72	0.797	0.231
	MRT[43]	ICCV'23	0.06	0.11	0.29	0.44	0.65	0.76	0.825	0.304
	Ours		**0.14**	**0.20**	**0.37**	**0.54**	**0.72**	**0.83**	**0.825**	0.312

Table S3: Table 1 in the main paper showed similar comparison with SOTA UDA methods on few dataset pairs. Here we show results for few more pairs.

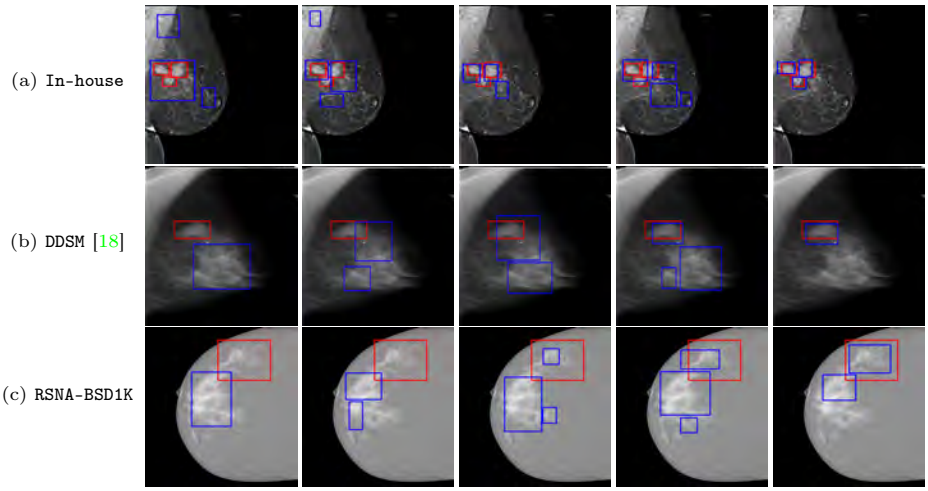

Fig. S1: Qualitative result comparison on in-house, DDSM, and RSNA-BSD1K datasets. Red boxes show the ground truth, and blue boxes show the predictions.