Supplementary Material

Lingrui Li[1,2], **Yanfeng Zhou**[1,2], **Ge Yang**[1,2] *
[1]Institute of Automation, Chinese Academy of Sciences
[2]School of Artifical Intelligence, University of Chinese Academy of Sciences
{lilingrui2021, zhouyanfeng2020, ge.yang}@ia.ac.cn

1. Results

Table S1 summarizes the results of the performance comparison of different competing robust methods and our method. All the methods perform better on adversarial samples. Our method achieves the best result overall.

2. Examples

Figure S1 shows the examples of (a) target fundus images and their (b) pseudo labels and (c) pseudo boundaries.

(a) Input Image   (b) Pseudo Label   (c) Pseudo Boundary
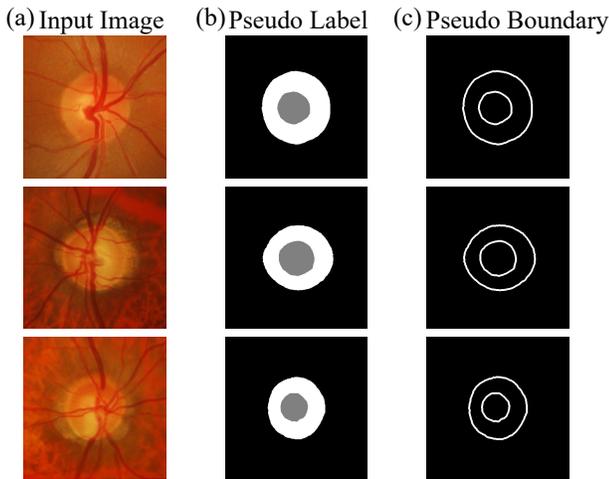
Figure S1. Examples of pseudo labels and boundaries of target images from the Drichti-GS dataset. Gray color region: OC; White: OD.

3. More Ablation Study

To test the effectiveness of pseudo boundaries, we performed two groups of experiments. In the first group of experiments, we performed ablation on both clean and adversarial samples for our method and find that adding boundary loss Lbl improves performance. The results are summarized in Table 3 and Table 4 of the original manuscript. Figure S2 shows representative examples in which using pseudo

boundaries to restrict boundary predictions generates more precise and coherent predictions, especially in e.g., low-contrast and small region.

Utilizing pseudo boundaries to enhance performance of domain adaptation is motivated by the observation that segmentation errors are more likely to occur near boundaries of image objects. Similar to pseudo labels, pseudo boundaries help self-supervised learning by making neural network models sensitive to edges of the image objects to be segmented, as shown in Figure S2 (b). By jointly optimizing pixel labeling and boundary accuracy, our method outperforms competing methods with no restrictions on boundary predictions.

We performed a second group of experiments in which we integrated our pseudo boundary scheme with the other competing methods. Ablation experiments on both adversarial and clean samples (Table S2 and S3.) show that using pseudo boundaries consistently improves their performance, further supporting effectiveness of pseudo boundaries.

(a) GT   (b) w/ L_{bl} (c) w/o L_{bl}
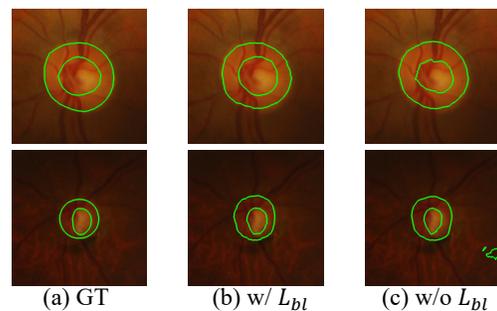
Figure S2. Comparison of segmentation results (b) with pseudo boundaries and (c) without pseudo boundaries versus (a) GT.

References

[1] Cheng Chen, Quande Liu, Yueming Jin, Qi Dou, and Pheng-Ann Heng. Source-free domain adaptive fundus image segmentation with denoised pseudo-labeling. In *International*

---

*Corresponding author.

Table S1. Quantitative comparison of different robust methods on the target domain datasets. Abbreviations: R: robust metrics; R.Dice: robust Dice; R.ASD: robust ASD; ST.S: standard source model; R.S.: robust source model.

Method	Optic disc segmentation				Optic cup segmentation			
	Dice	ASD	R.Dice	R.ASD	Dice	ASD	R.Dice	R.ASD
RIM-ONE-r3								
AdvEnt [4] robust	90.68	10.35	75.18	29.70	72.38	13.85	51.13	32.26
BEAL [6] robust	85.05	22.46	54.78	91.95	73.42	18.98	38.53	49.27
DPL [1] robust	87.29	11.95	55.55	44.27	75.51	11.16	40.52	45.33
TT-SFDA [3] robust	87.57	12.16	53.45	55.83	77.09	9.00	35.62	43.23
TENT [5] robust	84.66	17.22	72.11	28.05	76.44	11.10	60.03	29.25
OCDA [2] robust	87.92	13.70	65.88	48.59	76.93	10.83	52.70	51.05
SFDA-FSM [7] robust	88.45	11.50	75.99	25.29	70.50	18.64	62.82	_27.80_
Ours (Standard source)	_92.43_	_6.96_	74.35	35.58	_77.82_	_10.03_	55.91	45.40
Ours (Robust source)	91.39	8.07	**77.22**	_23.63_	75.59	11.13	**60.34**	50.11
Ours (Both)	**92.89**	**6.52**	_76.20_	**22.92**	**77.94**	10.07	_60.13_	**26.83**
Drishti-GS								
AdvEnt [4] robust	93.26	7.45	91.48	9.26	64.80	23.44	61.17	25.56
BEAL [6] robust	94.94	9.23	92.86	12.49	79.58	21.36	76.48	24.82
DPL [1] robust	95.40	4.70	95.02	5.57	83.55	11.37	77.47	15.82
TT-SFDA [3] robust	95.72	4.86	94.17	6.50	77.64	14.73	69.77	19.79
TENT [5] robust	94.53	6.67	94.20	6.77	80.16	13.48	75.02	17.02
OCDA [2] robust	95.87	5.08	92.36	9.35	77.08	15.59	63.62	25.13
SFDA-FSM [7] robust	95.47	5.00	80.74	28.34	78.85	14.37	67.70	16.39
Ours (Standard Source)	_96.01_	_4.70_	92.05	11.25	_83.71_	**10.91**	73.92	18.15
Ours (Robust Source)	95.67	5.09	**95.55**	**5.13**	82.87	11.55	_78.08_	_15.11_
Ours (Both)	**96.51**	**4.01**	_95.29_	_5.25_	83.56	_11.11_	**80.02**	**13.69**
Open Domain								
AdvEnt [4] robust	90.27	7.45	71.99	51.20	64.80	23.44	34.05	81.72
BEAL [6] robust	91.45	5.83	30.18	64.20	80.08	7.55	8.99	70.74
DPL [1] robust	89.04	16.64	80.43	22.75	79.18	15.91	68.36	32.26
TT-SFDA [3] robust	82.49	32.98	49.14	55.96	77.68	15.05	42.40	56.58
TENT [5] robust	89.42	17.98	79.92	34.55	_80.33_	10.27	71.15	26.81
OCDA [2] robust	89.73	18.79	58.49	55.09	80.08	11.53	47.73	49.73
SFDA-FSM [7] robust	91.37	16.93	79.87	44.19	75.82	8.98	32.27	35.93
Ours (Standard source)	91.54	6.87	88.51	17.68	79.78	7.44	46.46	37.46
Ours (Robust source)	_91.86_	_6.77_	_91.11_	_8.40_	**80.40**	_7.14_	**72.31**	_25.71_
Ours (Both)	**92.53**	**6.54**	**91.35**	**8.11**	80.31	**7.12**	_71.05_	**20.87**

Table S2. Ablation results with different losses on **adversarial** samples. (D ↑: %, ASD ↓: pixel.)

Method	Compound(C)				Open(O)		Avg.			
	RIM-ONE-r3		Drishti-GS		REFUGE val		C		C + O	
	D[%]	ASD	D[%]	ASD	D[%]	ASD	D[%]	ASD	D[%]	ASD
TT-SFUDA	53.22	61.21	77.83	18.14	46.21	54.13	65.52	39.68	59.08	44.49
TT-SFUDA+L_{bl}	56.01	42.69	77.97	18.12	46.06	51.73	**66.99**	**30.40**	**60.01**	**37.51**
TENT	53.71	51.70	81.75	16.04	51.76	50.55	67.73	33.87	62.41	39.43
TENT+L_{bl}	62.75	35.50	77.42	17.25	70.34	30.91	**70.08**	**26.37**	**70.17**	**27.89**
SFDA-FSM	63.94	34.12	80.24	16.65	58.65	37.59	72.09	25.39	67.61	29.46
SFDA-FSM+L_{bl}	65.17	32.96	81.11	15.08	61.70	35.76	**73.14**	**24.02**	**69.33**	**27.93**

Table S3. Ablation results with different losses on **clean** samples.

Method	Compound(C)				Open(O)		Avg.			
	RIM-ONE-r3		Drishti-GS		REFUGE val		C		C + O	
	D[%]	ASD	D[%]	ASD	D[%]	ASD	D[%]	ASD	D[%]	ASD
TENT	77.93	18.81	87.09	10.54	83.48	10.62	82.51	14.68	82.83	13.33
TENT+L_{bl}	81.81	10.88	82.05	10.26	81.93	10.57	81.97	10.59		
TT-SFUDA	80.81	13.68	87.95	9.50	80.23	22.78	84.38	11.59	83.00	15.32
TT-SFUDA+L_{bl}	82.54	12.80	86.91	9.32	80.34	20.60	**84.72**	**11.06**	**83.26**	**14.24**
SFDA-FSM	78.27	19.10	89.51	9.85	79.42	8.62	83.89	14.48	82.40	12.52
SFDA-FSM+L_{bl}	85.42	8.77	89.54	9.15	81.68	8.04	**87.48**	**8.96**	**85.54**	**8.65**

Conference on Medical Image Computing and Computer-Assisted Intervention, pages 225–235. Springer, 2021. 2

[2] Ziwei Liu, Zhongqi Miao, Xingang Pan, Xiaohang Zhan, Dahua Lin, Stella X Yu, and Boqing Gong. Open compound domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12406–12415, 2020. 2

[3] Vibashan VS, Jeya Maria Jose Valanarasu, and Vishal M Patel. Target and task specific source-free domain adaptive image segmentation. *arXiv preprint arXiv:2203.15792*, 2022. 2

[4] Tuan-Hung Vu, Himalaya Jain, Maxime Bucher, Matthieu Cord, and Patrick Pérez. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2517–2526, 2019. 2

[5] Dequan Wang, Evan Shelhamer, Shaoteng Liu, Bruno Olshausen, and Trevor Darrell. Tent: Fully test-time adaptation by entropy minimization. In *International Conference on Learning Representations*, 2020. 2

[6] Shujun Wang, Lequan Yu, Kang Li, Xin Yang, Chi-Wing Fu, and Pheng-Ann Heng. Boundary and entropy-driven adversarial learning for fundus image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 102–110. Springer, 2019. 2

[7] Chen Yang, Xiaoqing Guo, Zhen Chen, and Yixuan Yuan. Source free domain adaptation for medical image segmentation with fourier style mining. *Medical Image Analysis*, 79:102457, 2022. 2