Dual Knowledge-Aware Guidance for Source-Free Domain Adaptive Fundus Image Segmentation

Yu Chen[1,2], Hailing Wang[1,2], Chunwei Wu[3], and Guitao Cao[1,2(✉)]

[1] Shanghai Key Laboratory of Trustworthy Computing, East China Normal University, Shanghai 200062, China
[2] MoE Engineering Research Center of SW/HW Co-design Technology and Application, East China Normal University, Shanghai 200062, China
gtcao@sei.ecnu.edu.cn
[3] Ant Group, Hangzhou 310023, China

Abstract. Source-free domain adaptation (SFDA), where only a pre-trained source model is available to adapt to the target domain, has gained widespread application in the medical field. Most existing methods overlook low-quality pseudo-labels, i.e., pseudo-labels with boundary semantic confusion, when learning target domain-specific knowledge, leading to the loss of crucial boundary information. Furthermore, focusing solely on the specific knowledge can drive the model shifts in an uncontrollable direction, resulting in model degradation. To address these issues, we propose Dual Knowledge-aware Guidance (DKG), a novel SFDA method that integrates domain-specific knowledge with domain-invariant knowledge to improve transfer performance. Specifically, the pseudo-label calibration scheme is proposed to reduce semantic bias in high-uncertainty pixels, preserving the boundary information of target domain-specific knowledge. To ensure stable training, we propose a domain-invariant knowledge-based loss strategy, leveraging a confidence-guided mechanism and a consistency constraint. Additionally, we also introduce a dynamic balancing loss to address class imbalance. Extensive experiments on cross-domain fundus image segmentation show that DKG achieves state-of-the-art performance. Code is available at https://github.com/Hanshuqian/DKG.

Keywords: Source-free domain adaptation · Fundus image · Pseudo-label calibration · Domain-invariant knowledge.

1 Introduction

Medical image segmentation is a critical visual task in computer-aided diagnosis. While deep neural networks have made remarkable advancements in this field, dealing with domain shifts caused by different scanning devices or hospitals remains a significant challenge [25]. Unsupervised domain adaptation (UDA) strives to alleviate domain shift problem. However, typical UDA methods [30,9,27] require coordinated access to source and target domains, which can be limited by data privacy and security issues, especially in medical scenes. In order to

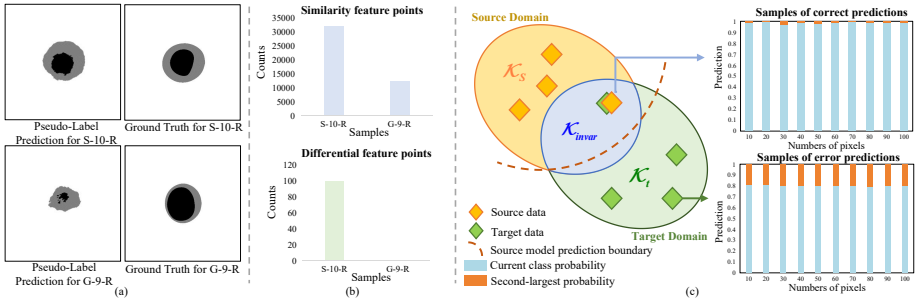

Fig. 1. Analysis of domain-invariant knowledge. (a) Examples of inconsistent prediction accuracy using the plain pseudo-labeling method. (b) Numbers of similarity and differential feature points. (c) Visualization of two probability values in the pixel-level random sampling average probability distribution.

improve the practical application of the models, source-free domain adaptation (SFDA) [6,11,13,14,18,20] emerged. The core idea is to adapt a pre-trained source model to the unlabeled target domain without using any source data.

Existing SFDA methods for image segmentation can be divided into two mainstream types: generation-based and self-training methods. Generation-based methods [28,22] aim to produce source-like samples or distributions to compensate for unavailable source data. However, the quality of these proxies poses a new challenge. In contrast to the previous method, self-training methods [3,10,29,5] directly utilize model predictions on unlabeled target data for self-optimization, making them more practical and flexible in actual applications. An intuitive strategy is to generate pseudo-labels for the target domain using the source model. Due to the domain shift, pseudo-labels inevitably contain noise. To improve robustness, [3] proposes a dual denoising scheme at pixel and class levels via uncertainty estimation and prototype evaluation. [10] refines pseudo-labels by addressing context inconsistency, while [29] introduces adaptive thresholding through multi-class negative learning to select pseudo-labels.

Current SFDA methods typically regard high-quality pseudo-labels, such as those with distinguishable semantics, as critical for transfer performance. However, exclusive reliance on these labels can lead to a loss of semantic boundary information during training, impeding the capture of discriminative class features. Additionally, as shown in Fig. 1(a), the impact of domain shift on different target samples varies significantly, and this inconsistency exacerbates training instability, potentially even leading to model degradation. Furthermore, Fig. 1(b) demonstrates that high-accuracy samples preserve more domain-shared features, facilitating the learning of domain-specific features. Moreover, Fig. 1(c) reveals the correlation between prediction confidence and domain-invariant knowledge $\mathcal{K}_{invar}$ (the overlapping region between source domain-specific knowledge $\mathcal{K}_s$ and target domain-specific knowledge $\mathcal{K}_t$) through pixel-level random sampling on target data using a non-adapted source model.

To address these challenges, we propose the Dual Knowledge-aware Guidance (DKG) method. Firstly, to effectively capture semantic boundary information of pseudo-labels, we propose an Anchor Point-Driven Pseudo-Label Calibration scheme. Leveraging semantic distribution consistency between reference regions (anchor points) and high-uncertainty pixels in high-dimensional feature space, the scheme calibrates the semantics of pseudo-labels with high-uncertainty, which strives to mine domain-specific knowledge $\mathcal{K}_t$ during target domain adaptation. Additionally, global class prototype representation is employed to ensure anchor points' reliability. Secondly, to enhance training stability, we propose a loss strategy based on domain-invariant knowledge $\mathcal{K}_{invar}$. This strategy leverages a confidence-guided mechanism and is further constrained by the consistency between the uncertainty map and the class prototype distance. Finally, to balance foreground and background during target domain-specific knowledge learning, we design an adaptive balancing loss to adjust class contributions. Experiments on cross-domain fundus image segmentation demonstrate that our method achieves SOTA performance.

2 Method

In the setting of SFDA, we are given a model $f_s : \mathcal{X}_s \longrightarrow \mathcal{Y}_s$ trained on the source domain $D_s = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$, along with an unlabeled dataset $\{x_i^t\}_{i=1}^{N_t}$ from target domain D_t. The goal of SFDA is to adapt the pre-trained model f_s to the target domain D_t and obtain an adapted model $f_t : \mathcal{X}_t \longrightarrow \mathcal{Y}_t$ with high accuracy. In this paper, we consider the fundus image segmentation, $x_i \in \mathbb{R}^{H\times W\times 3}$ and $y_i \in \{0,1\}^{H\times W\times C}$, where C representing the number of classes and $C = 2$ as there are two segmentation targets: optic cup and optic disc.

Fig. 2 illustrates our Dual Knowledge-aware Guidance (DKG) method. In this section, we first introduce an Anchor Point-Driven Pseudo-Label Calibration scheme that effectively sharpens the semantic discriminative boundaries. Then, we propose a domain-invariant knowledge-based loss strategy, followed by a balancing loss designed for learning specific knowledge.

2.1 Anchor Point-Driven Pseudo-Label Calibration

Ensuring the accurate learning of domain-specific knowledge is key to improving model transfer performance. However, significant domain distribution differences often lead to semantic boundary bias in pseudo-labels. Different from the [16], we focus on pixel-level semantic confidence variations. Leveraging the contextual consistency of semantic information in the feature space, we propose the Anchor Point-Driven Pseudo-Label Calibration (ADPC) scheme (see the blue region in Fig. 2). ADPC effectively calibrates the pseudo-label boundary information, enhancing the global semantics of domain-specific knowledge.

Pseudo-Label Calibration. Epistemic uncertainty aims to quantify the reliability of predictions [3]. In this work, we model it using Monte Carlo Dropout

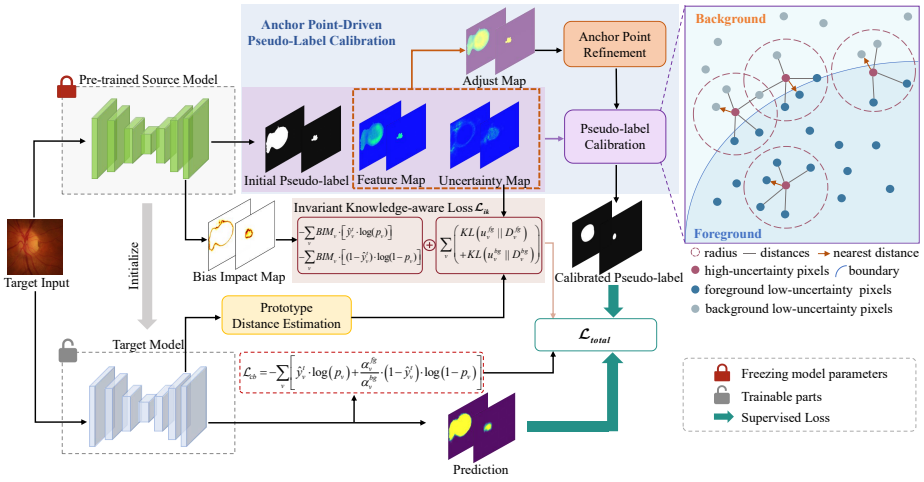

Fig. 2. Overview of our proposed DKG. For a target input, the source model generates an initial pseudo-label, and low-uncertainty pixels are refined using uncertainty map and feature map, serving as anchor points for calibration (illustrated on the right). Guided by the bias impact map, $\mathcal{L}_{ik}$ is computed by aligning the uncertainty map with the class prototype distance. $\mathcal{L}_{cb}$ is introduced to address class imbalance. The total loss integrates both losses.

to obtain varying-quality pseudo-label regions. For a given target image x^t, K random forward passes using f_s, yielding prediction outputs $p_k = f_s(x^t), k = 1, \ldots, K$. The initial pseudo-label and uncertainty map are defined as:

$$p = \frac{1}{K}\sum_{k=1}^{K} p_k, \quad \hat{y}^t = \mathbb{1}[p > \gamma], \quad u = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(p_k - p)^2}, \tag{1}$$

where $\mathbb{1}$ is the indicator function, $\gamma \in (0, 1)$ is a probability threshold for generating binary pseudo-labels. Previous methods commonly use the low-uncertainty mask $m^{low} = \mathbb{1}[u < \eta]$ (η is the uncertainty threshold) to obtain pseudo-labels for supervised learning. We further introduce the high-uncertainty mask $m^{high} = \mathbb{1}[u \geq \eta]$ and integrate both masks to preserve complete semantic boundary information. Specifically, for each high-uncertainty pixel h, we search for the nearest low-uncertainty pixel within a radius r in the feature space, where the feature map F is obtained from the layer before the last convolution. Using the low-uncertainty pixels j as anchor points, we calibrate the semantic information of the high-uncertainty pixel:

$$\hat{y}_h^t = \hat{y}_j^t : \underset{d(h,j)\leq r}{argmin}\left(1 - \frac{F_h \cdot F_j}{|F_h| \cdot |F_j|}\right). \tag{2}$$

Anchor Point Refinement. Low-uncertainty pixels as anchor points are prone to being influenced by feature space error propagation. Directly using unopti-

mized anchor points can lead to confirmation bias [17]. To address this, we propose a global representation-based approach to reduce semantic estimation bias in low-uncertainty pixels. Specifically, we define the low-uncertainty masks for foreground and background as $m^{fg} = \mathbb{1}[\hat{y}^t = 1]\mathbb{1}[u < \eta]$ and $m^{bg} = \mathbb{1}[\hat{y}^t = 0]\mathbb{1}[u < \eta]$, respectively. Then, we compute the foreground class prototype g^{fg} from each pixel v, the global class prototype G, and the adjust map A as follows:

$$g^{fg} = \frac{\sum_v F_v \cdot m_v^{fg} \cdot p_v}{\sum_v m_v^{fg} \cdot p_v}, \tag{3}$$

$$G = \frac{1}{N_t} \sum_{i=1}^{N_t} g_i^{fg}, \quad S = \frac{F \cdot G}{\|F\| \|G\|}, \quad A = S \odot \mathbb{1}\left[\hat{y}^t = 1\right]. \tag{4}$$

We retain the highly similar regions in A, and the low-uncertainty mask is refined as $m^{relow} = m^{fg} \odot \mathbb{1}\left[A > \delta\right] + m^{bg}$.

2.2 Invariant Knowledge-Aware Loss

As shown in Fig. 1(c), domain-invariant knowledge $\mathcal{K}_{invar}$ (shared knowledge across domains), target domain-specific knowledge $\mathcal{K}_t$ and source domain-specific knowledge $\mathcal{K}_s$ form a complete knowledge framework. Overlooking the extraction of $\mathcal{K}_{invar}$ during training may limit the model's cross-domain generalization ability, leading to under-transfer. To tackle this problem, we explore $\mathcal{K}_{invar}$ from two perspectives. The learning process of the pre-trained model on the target domain can be expressed as:

$$\hat{y}^t = f_s\left(x^t\right) = \underbrace{f_t\left(x^t\right) - \mathcal{K}_t\left(x^t\right)}_{\mathcal{K}_{invar}(x^t)} + \mathcal{K}_s\left(x^t\right) \approx y^t + \triangle\left(\mathcal{K}_s, \mathcal{K}_t\right), \tag{5}$$

where y^t denotes the expected label. From Eq. 5, it is evident that the knowledge bias $\triangle\left(\mathcal{K}_s, \mathcal{K}_t\right)$ undermines the pre-trained model's ability to represent target domain samples, reducing prediction confidence. Statistical results in Fig. 1(c) show that the model exhibits higher decision confidence on correctly predicted pixels than incorrect ones, as indicated by the probability ratio between the second-largest and the current class. This suggests a positive correlation between the ratio and knowledge bias. To quantify the impact of knowledge bias, we introduce the Bias Impact Map (BIM) to measure its effect on each pixel v:

$$BIM_v = 1 - \frac{\max\limits_{c \in C, c \neq \hat{y}_v^t} p_v^c}{\max\limits_{c \in C} p_v^c}. \tag{6}$$

The BIM-based loss is formulated as:

$$\mathcal{L}_{bim} = -\sum_v BIM_v \cdot \left[\hat{y}_v^t \cdot \log(p_v) + (1 - \hat{y}_v^t) \cdot \log(1 - p_v)\right]. \tag{7}$$

Eq. 7 guides the model to focus on domain-invariant knowledge through confidence but does not explicitly align knowledge structures between source and target domains, potentially leading to insufficient extraction. Inspired by this, we introduce Kullback-Leibler divergence to match prediction uncertainty with the prototype distance, effectively capturing consistent information. Specifically, for a target sample x^t and its uncertainty map u, we compute class prototypes g^{fg} (foreground) and g^{bg} (background) using Eq. 3. The feature distance between each feature vector for pixel v to the two prototypes is then calculated as $d_v^\omega = \|F_v - g^\omega\|_2$, where $\omega \in \{fg, bg\}$. To ensure robustness, we focus only on low-uncertainty regions:

$$u_v^\omega = u_v \cdot m_v^\omega, \quad D_v^\omega = d_v^\omega \cdot m_v^\omega. \tag{8}$$

Eq. 7 is regularized to obtain the invariant knowledge-aware loss:

$$\mathcal{L}_{ik} = \mathcal{L}_{bim} + \sum_v \left(KL\left(u_v^{fg} || D_v^{fg}\right) + KL\left(u_v^{bg} || D_v^{bg}\right)\right). \tag{9}$$

2.3 Class-Wise Balancing Loss

Balancing the contribution of different classes remains a significant challenge in medical image segmentation, as biased predictions toward a single dominant class often result in trivial solutions [2]. Previous researchers typically use class-weighted optimization based on pixel counts [23,7]. However, this approach heavily depends on pseudo-label accuracy and faces challenges in adapting to target domain-specific knowledge early in training. Therefore, we introduce a dynamic balancing loss based on batch-level knowledge.

Specifically, we calculate the average loss of foreground and background for N_b samples in the current batch:

$$\alpha^{fg} = \frac{\sum_{i=1}^{N_b} \mathcal{L}_{bce}^i \cdot \mathbb{1}\left[\hat{y}_i = 1\right]}{\sum_{i=1}^{N_b} \mathbb{1}\left[\hat{y}_i = 1\right]}, \quad \alpha^{bg} = \frac{\sum_{i=1}^{N_b} \mathcal{L}_{bce}^i \cdot \mathbb{1}\left[\hat{y}_i = 0\right]}{\sum_{i=1}^{N_b} \mathbb{1}\left[\hat{y}_i = 0\right]}, \tag{10}$$

where $\mathcal{L}_{bce} = -\sum_v \left[\hat{y}_v^t \cdot \log\left(p_v\right) + \left(1 - \hat{y}_v^t\right) \cdot \log\left(1 - p_v\right)\right]$. Our balancing loss is given as:

$$\mathcal{L}_{cb} = -\sum_v \left[\hat{y}_v^t \cdot \log\left(p_v\right) + \frac{\alpha_v^{fg}}{\alpha_v^{bg}} \cdot \left(1 - \hat{y}_v^t\right) \cdot \log\left(1 - p_v\right)\right]. \tag{11}$$

2.4 Total Loss Function

To harmonize target domain-specific knowledge with domain-invariant knowledge, the total loss integrates both strategies:

$$\mathcal{L}_{total} = \sum_v BIM_v \cdot \mathcal{L}_{cb} + \sum_v \left(KL\left(u_v^{fg} || D_v^{fg}\right) + KL\left(u_v^{bg} || D_v^{bg}\right)\right). \tag{12}$$

Table 1. Comparison with other methods on two datasets. S-F refers to source-free. ± refers to the standard deviation across samples.

Methods	S-F	Optic Disc Segmentation		Optic Cup Segmentation	
		Dice[%]↑	ASD[pixel]↓	Dice[%]↑	ASD[pixel]↓
RIM-ONE-r3					
W/o DA [3]		83.18±6.46	24.15±15.58	74.51±16.40	14.44±11.27
Oracle [25]		96.80	–	85.60	–
BEAL [25]	×	89.80	–	81.00	–
AdvEnt [24]	×	89.73±3.66	9.84±3.86	77.99±21.08	7.57±4.24
SRDA [1]	✓	89.37±2.70	9.91±2.45	77.61±13.58	10.15±5.75
DAE [12]	✓	89.08±3.32	11.63±6.84	79.01±12.82	10.31±8.45
DPL [3]	✓	90.13±3.06	9.43±3.46	79.78±11.05	9.01±5.59
CPR [10]	✓	91.51±3.20	7.75±3.09	77.11±18.02	9.12±4.30
EOAPNet [5]	✓	92.61±3.13	6.67±2.91	74.59±25.64	8.74±5.34
DKG(Ours)	✓	**92.68±3.60**	**6.50±2.84**	**83.35±15.63**	**6.90±4.14**
Drishti-GS					
W/o DA [3]		93.84±2.91	9.05±7.50	83.36±11.95	11.39±6.30
Oracle [25]		97.40	–	90.10	–
BEAL [25]	×	96.10	–	86.20	–
AdvEnt [24]	×	96.16±1.65	4.36±1.83	82.75±11.08	11.36±7.22
SRDA [1]	✓	96.22±1.30	4.88±3.47	80.67±11.78	13.12±6.48
DAE [12]	✓	94.04±2.85	8.79±7.45	83.11±11.89	11.56±6.32
DPL [3]	✓	96.39±1.33	4.08±1.49	83.53±17.80	11.39±10.18
CPR [10]	✓	96.22±2.42	4.91±4.86	85.06±12.84	9.92±5.76
EOAPNet [5]	✓	96.46±1.66	4.00±1.83	79.75±10.31	13.95±6.55
DKG(Ours)	✓	**96.70±1.46**	**3.72±1.55**	**87.12±12.27**	**8.33±5.18**

3 Experiments

Dataset. We evaluate our method on three widely-used fundus segmentation datasets from different clinical centers: the REFUGE challenge training set [15] as the source domain, the RIM-ONE-r3 [8] and Drishti-GS [21] datasets as target domains. The source domain contains 400 images with labels, and the two target domain data are divided into 99/60 and 50/51 for training/testing. We apply the same preprocessing method as [3] for a fair comparison.

Implementation Details and Evaluation Metrics. We use DeepLabv3+ [4] with a MobileNetV2 [19] backbone as our segmentation model, following the previous work [3,25,26]. The threshold γ for pseudo-label selection is set to 0.75, as in [25], with an uncertainty threshold $\eta = 0.02$ by visually inspecting and radius $r = 5$ in Eq. 2. The similarity threshold δ for the adjust map A is empirically set to 0.7. We use the Adam optimizer with momentum coefficients of 0.9 and 0.99. During target domain adaptation, the learning rate is 2e-3, and the batch size is 8. Our method avoids model degradation, enabling stable training for 20 epochs. Our framework is implemented in PyTorch, using an NVIDIA RTX A40

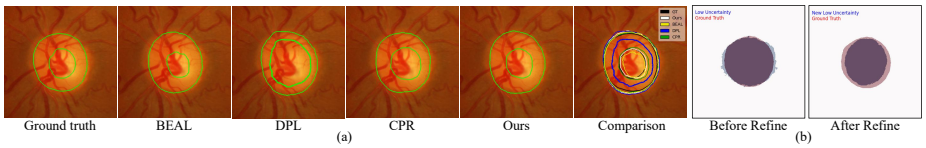

Fig. 3. Drishti-GS as the target domain: (a) Qualitative comparison across different methods. (b) An example of anchor point change using the proposed refinement scheme.

Table 2. Results of ablation study on the REFUGE to RIM-ONE-r3 adaptation.

ADPC		$\mathcal{L}_{total}$		Avg. Dice[%]↑	Avg. ASD[pixel]↓
w/o refine	w/ refine	$\mathcal{L}_{ik}$	$\mathcal{L}_{cb}$		
×	×	×	×	82.96	9.97
✓	×	×	×	86.07	7.80
✓	✓	×	×	86.80	7.53
✓	✓	×	✓	87.30	6.95
✓	✓	✓	×	87.33	7.04
×	×	✓	✓	84.61	7.47
✓	✓	✓	✓	**88.02**	**6.70**

GPU. To evaluate our proposed method, we use two standard metrics, including Dice Coefficient (Dice) and Average Surface Distance (ASD).

Comparison with State-of-the-Arts. We compare our method with several state-of-the-art domain adaptation approaches, including the baseline results without domain adaptation ("Oracle", as a lower bound) and fully supervised learning results ("W/o DA", as an upper bound). As shown in Table 1, our method outperforms existing SFDA methods across all metrics and even surpasses traditional UDA approaches. Notably, in the challenging optic cup segmentation task, our method achieves Dice improvements of 3.57% and 2.06% on two datasets. Visualization results are presented in Fig. 3(a).

Ablation Study. We evaluate the contributions of each component in our method using RIM-ONE-r3. As shown in Table 2, pseudo-label calibration significantly improves segmentation performance, highlighting its effectiveness in correcting low-quality pseudo-label semantics. The comparison between w/o refine and w/ refine emphasizes the importance of anchor point refinement, as visualized in Fig. 3(b). Incorporating $\mathcal{L}_{ik}$ leverages domain-invariant knowledge, consistently boosting all metrics compared to using ADPC alone. In addition, $\mathcal{L}_{cb}$ reduces ASD by addressing class imbalance in domain-specific learning, improving boundary segmentation accuracy. Furthermore, compared to using the two losses individually, our total loss function represents an effective combination strategy. These results demonstrate the effectiveness of each DKG component.

4 Conclusion

This paper presents a Dual Knowledge-aware Guidance method for SFDA. For the semantic boundary confusion in pseudo-labels, we propose an Anchor Point-Driven Pseudo-Label Calibration scheme that ensures full capture of target domain-specific knowledge. For stabilizing model training, we propose a domain-invariant knowledge-aware loss strategy to keep cross-domain feature distribution alignment. Additionally, we employ a dynamic balancing loss to address the class imbalance in target data. Experimental results on cross-domain fundus image segmentation demonstrate that DKG achieves state-of-the-art performance.

Acknowledgments. This work was supported by the National Natural Science Foundation of China under Grant 61871186 and 61771322.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bateson, M., Kervadec, H., Dolz, J., Lombaert, H., Ayed, I.B.: Source-relaxed domain adaptation for image segmentation. In: MICCAI (1). Lecture Notes in Computer Science, vol. 12261, pp. 490–499. Springer (2020)
2. Bateson, M., Kervadec, H., Dolz, J., Lombaert, H., Ayed, I.B.: Source-free domain adaptation for image segmentation. Medical Image Anal. **82**, 102617 (2022)
3. Chen, C., Liu, Q., Jin, Y., Dou, Q., Heng, P.: Source-free domain adaptive fundus image segmentation with denoised pseudo-labeling. In: MICCAI (5). Lecture Notes in Computer Science, vol. 12905, pp. 225–235. Springer (2021)
4. Chen, L., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: ECCV (7). Lecture Notes in Computer Science, vol. 11211, pp. 833–851. Springer (2018)
5. Chen, Y., Li, J., Yu, H., Qi, L., Li, Y.: Source-free unsupervised domain adaptation fundus image segmentation via entropy optimization and anatomical priors. Procedia Computer Science **250**, 182–187 (2024)
6. Ding, N., Xu, Y., Tang, Y., Xu, C., Wang, Y., Tao, D.: Source-free domain adaptation via distribution estimation. In: CVPR. pp. 7202–7212. IEEE (2022)
7. Fidon, L., Li, W., García-Peraza-Herrera, L.C., Ekanayake, J., Kitchen, N., Ourselin, S., Vercauteren, T.: Generalised wasserstein dice score for imbalanced multi-class segmentation using holistic convolutional networks. In: BrainLes@MICCAI. Lecture Notes in Computer Science, vol. 10670, pp. 64–76. Springer (2017)
8. Fumero, F., Alayón, S., Sánchez, J.L., Sigut, J.F., González-Hernández, M.: RIM-ONE: an open retinal image database for optic nerve evaluation. In: CBMS. pp. 1–6. IEEE Computer Society (2011)
9. Hoyer, L., Dai, D., Wang, H., Gool, L.V.: MIC: masked image consistency for context-enhanced domain adaptation. In: CVPR. pp. 11721–11732. IEEE (2023)
10. Huai, Z., Ding, X., Li, Y., Li, X.: Context-aware pseudo-label refinement for source-free domain adaptive fundus image segmentation. In: MICCAI (7). Lecture Notes in Computer Science, vol. 14226, pp. 618–628. Springer (2023)

11. Jing, M., Zhen, X., Li, J., Snoek, C.: Variational model perturbation for source-free domain adaptation. In: NeurIPS (2022)
12. Karani, N., Erdil, E., Chaitanya, K., Konukoglu, E.: Test-time adaptable neural networks for robust medical image segmentation. Medical Image Anal. **68**, 101907 (2021)
13. Li, R., Jiao, Q., Cao, W., Wong, H., Wu, S.: Model adaptation: Unsupervised domain adaptation without source data. In: CVPR. pp. 9638–9647. Computer Vision Foundation / IEEE (2020)
14. Liang, J., Hu, D., Feng, J.: Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation. In: ICML. Proceedings of Machine Learning Research, vol. 119, pp. 6028–6039. PMLR (2020)
15. Orlando, J.I., Fu, H., Breda, J.B., van Keer, K., Bathula, D.R., Diaz-Pinto, A., Fang, R., Heng, P., Kim, J., Lee, J., Lee, J., Li, X., Liu, P., Lu, S., Murugesan, B., Naranjo, V., Phaye, S.S.R., Shankaranarayana, S.M., Bogunovic, H.: REFUGE challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. Medical Image Anal. **59** (2020)
16. Ren, Q., Lu, S., Mao, Q., Dong, M.: Exploring prototype-anchor contrast for semantic segmentation. IEEE Trans. Circuits Syst. Video Technol. **34**(8), 7106–7120 (2024)
17. Rizve, M.N., Duarte, K., Rawat, Y.S., Shah, M.: In defense of pseudo-labeling: An uncertainty-aware pseudo-label selection framework for semi-supervised learning. In: ICLR. OpenReview.net (2021)
18. Roy, S., Trapp, M., Pilzer, A., Kannala, J., Sebe, N., Ricci, E., Solin, A.: Uncertainty-guided source-free domain adaptation. In: ECCV (25). Lecture Notes in Computer Science, vol. 13685, pp. 537–555. Springer (2022)
19. Sandler, M., Howard, A.G., Zhu, M., Zhmoginov, A., Chen, L.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: CVPR. pp. 4510–4520. Computer Vision Foundation / IEEE Computer Society (2018)
20. Sivaprasad, P.T., Fleuret, F.: Uncertainty reduction for model adaptation in semantic segmentation. In: CVPR. pp. 9613–9623. Computer Vision Foundation / IEEE (2021)
21. Sivaswamy, J., Krishnadas, S., Chakravarty, A., Joshi, G., Tabish, A.S., et al.: A comprehensive retinal image dataset for the assessment of glaucoma from the optic nerve head analysis. JSM Biomedical Imaging Data Papers **2**(1), 1004 (2015)
22. Stan, S., Rostami, M.: Unsupervised model adaptation for continual semantic segmentation. In: AAAI. pp. 2593–2601. AAAI Press (2021)
23. Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Cardoso, M.J.: Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: DLMIA/ML-CDS@MICCAI. Lecture Notes in Computer Science, vol. 10553, pp. 240–248. Springer (2017)
24. Vu, T., Jain, H., Bucher, M., Cord, M., Pérez, P.: ADVENT: adversarial entropy minimization for domain adaptation in semantic segmentation. In: CVPR. pp. 2517–2526. Computer Vision Foundation / IEEE (2019)
25. Wang, S., Yu, L., Li, K., Yang, X., Fu, C., Heng, P.: Boundary and entropy-driven adversarial learning for fundus image segmentation. In: MICCAI (1). Lecture Notes in Computer Science, vol. 11764, pp. 102–110. Springer (2019)
26. Xu, Z., Lu, D., Wang, Y., Luo, J., Wei, D., Zheng, Y., Tong, R.K.: Denoising for relaxing: Unsupervised domain adaptive fundus image segmentation without source data. In: MICCAI (5). Lecture Notes in Computer Science, vol. 13435, pp. 214–224. Springer (2022)

27. Xu, Z., Gong, H., Wan, X., Li, H.: ASC: appearance and structure consistency for unsupervised domain adaptation in fetal brain MRI segmentation. In: MICCAI (7). Lecture Notes in Computer Science, vol. 14226, pp. 325–335. Springer (2023)
28. Yang, C., Guo, X., Chen, Z., Yuan, Y.: Source free domain adaptation for medical image segmentation with fourier style mining. Medical Image Anal. **79**, 102457 (2022)
29. Yang, C., Kuo, Y., Hsu, C.: Source free domain adaptation for semantic segmentation via distribution transfer and adaptive class-balanced self-training. In: ICME. pp. 1–6. IEEE (2022)
30. Zhang, Z., Chen, M., Xiao, S., Peng, L., Li, H., Lin, B., Li, P., Wang, W., Wu, B., Cai, D.: Pseudo label refinery for unsupervised domain adaptation on cross-dataset 3d object detection. In: CVPR. pp. 15291–15300. IEEE (2024)