

Advancing UWF-SLO Vessel Segmentation with Source-Free Active Domain Adaptation and a Novel Multi-Center Dataset

Hongqiu Wang^{1*}, Xiangde Luo^{2*}, Wu Chen², Qingqing Tang³, Mei Xin⁴, Qiong Wang⁵, and Lei Zhu^{1,6}(✉)

¹ Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China
leizhu@ust.hk

² University of Electronic Science and Technology of China, Chengdu, China

³ West China Hospital, Sichuan University, Chengdu, China

⁴ Chengdu First People's Hospital

⁵ Guangdong Provincial Key Laboratory of Computer Vision and Virtual Reality Technology, Shenzhen Institute of Advanced Technology

⁶ The Hong Kong University of Science and Technology

Abstract. Accurate vessel segmentation in Ultra-Wide-Field Scanning Laser Ophthalmoscopy (UWF-SLO) images is crucial for diagnosing retinal diseases. Although recent techniques have shown encouraging outcomes in vessel segmentation, models trained on one medical dataset often underperform on others due to domain shifts. Meanwhile, manually labeling high-resolution UWF-SLO images is an extremely challenging, time-consuming and expensive task. In response, this study introduces a pioneering framework that leverages a patch-based active domain adaptation approach. By actively recommending a few valuable image patches by the devised Cascade Uncertainty-Predominance (CUP) selection strategy for labeling and model-finetuning, our method significantly improves the accuracy of UWF-SLO vessel segmentation across diverse medical centers. In addition, we annotate and construct the first Multi-center UWF-SLO Vessel Segmentation (MU-VS) dataset to promote this topic research, comprising data from multiple institutions. This dataset serves as a valuable resource for cross-center evaluation, verifying the effectiveness and robustness of our approach. Experimental results demonstrate that our approach surpasses existing domain adaptation and active learning methods, considerably reducing the gap between the Upper and Lower bounds with minimal annotations, highlighting our method's practical clinical value. We will release our dataset and code to facilitate relevant research (<https://github.com/whq-xxh/SFADA-UWF-SLO>).

Keywords: Vessel segmentation · Ultra-Wide-Field · source free · active domain adaptation · multi-center dataset.

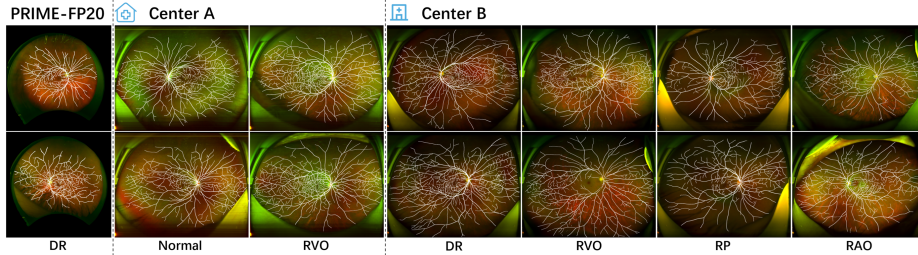
1 Introduction

Accurate segmentation of retinal vessels in fundus images is critical in aiding ophthalmologists with quantitative analysis and treatment [3,11]. For instance,

H. Wang and X. Luo contributed equally to this work.

Table 1. Quantitative analysis of the existing datasets, including categories and image resolution. Center A: Hospital A. Center B: Hospital B.

Dataset	Amount of data	Categories	Resolution	Public available
In-house [9]	65	Normal, VO	3900×3072	no
PRIME-FP20 [2]	15	DR	4000×4000	yes
Center A (Ours)	30	Normal, RVO	3900×3072	yes
Center B (Ours)	30	DR, RVO, RP, RAO, CSC	3900×3072	yes

**Fig. 1.** Several data visualization examples from the proposed Multi-center UWF-SLO Vessel Segmentation (MU-VS) dataset, illustrate that various centers encompass different disease categories.

Retinal Vein Occlusion (RVO) is identified by increased retinal vessel tortuosity, enlarged vessel caliber, and retinal non-perfusion [16]. A variety of deep learning models for automated vessel segmentation have emerged, showing promising results [4,27,28,29]. These segmentation models are mainly adapted to Narrow Field (NF) Fundus Photography (FP), since NF FP is the most common format and modality in clinical practice, and previous relevant segmentation datasets are also mainly based on NF FP [6,18].

Recently, Ultra-Wide-Field Scanning Laser Ophthalmoscopy (UWF-SLO) imaging has gained popularity due to its ability to provide extensive retinal coverage and superior imaging of peripheral lesions over NF FP, thereby enhancing diagnostic precision [21]. Generally, UWF-SLO images provide an expansive 200° field-of-view (FOV), far exceeding the typical 30°-50° FOV of NF FP. This broader view grants ophthalmologists access to more comprehensive information for more accurate diagnoses [2,12]. There are also some efforts in the field. For example, Li *et al.* proposed a weakly-supervised iterative learning method and the PRIME-P20 dataset to segment vessels in UWF images[2]. Qiu *et al.* introduced a dual-stream super-resolution network for this task[15].

Although there has been a notable improvement in UWF vessel segmentation approaches, the above models are all developed within the single-center setting, lacking multi-center, cross-domain studies. In real-world clinical scenarios, domain shifts commonly occur among datasets [5,31] because of variations in imaging equipment and patient populations, potentially leading to suboptimal performance in new domains [13,26]. The simplest approach involves fully supervised training on target data, yet annotating high-resolution UWF-SLO images is extremely expensive, requiring approximately 18 hours of an expert’s time to annotate a single image [14]. Therefore, unsupervised domain adaptation (UDA) techniques are widely explored, designed to reduce the domain discrepancy between the labeled source and the unlabeled target domain [8]. Although

UDA approaches yield better outcomes, their performance still significantly falls short of that achieved by fully supervised models [8,10,26]. Moreover, accessing source medical datasets raises privacy and security concerns [30].

To alleviate these above issues, we propose a novel patch-based Source-Free Active Domain Adaptation (SFADA) method for advancing UWF-SLO cross-center vessel segmentation. Our approach offers three advantages: First, it eliminates the need to access source domain data, thereby enhancing data security and privacy protection. Second, we introduce the Cascade Uncertainty-Predominance (CUP) selection strategy, which efficiently identifies a small subset of valuable image patches for annotation, substantially reducing the annotation burden. Lastly, by integrating our method with a minimal number of patch annotations, we can significantly boost the model’s performance. Meanwhile, by integrating the existing dataset and our newly collected and labeled datasets from two distinct centers (as detailed in Table 1 and Fig. 1), we construct the first Multi-center UWF-SLO Vessel Segmentation (MU-VS) dataset to explore the cross-center segmentation study. The main contributions are summarized as follows:

- To our knowledge, this marks a pioneering exploration on an essential application of cross-center vessel segmentation using UWF-SLO, and we propose a patch-based SFADA framework to enhance segmentation performance.
- We design the Cascade Uncertainty Predominance (CUP) selection strategy to select a small number of patches with high uncertainty and dominance to recommend for manual annotation.
- We establish the first multi-center UWF-SLO vessel segmentation dataset consisting of 60 UWF-SLO images from two hospitals, named MU-VS, to support relevant studies.
- Experimental results show that our method significantly surpasses other state-of-the-art domain adaptation and active learning methods, effectively enhancing segmentation accuracy.

2 Methodology

2.1 Problem setting

The goal of medical image segmentation is to construct a model \mathcal{M} that links an image sample x from the space X to its predictive label y within the space Y . In the SFADA setting, direct access to the source dataset and its annotations (x^s, y^s) is avoided, thereby safeguarding data privacy and security. Instead, we employ a model \mathcal{M}^s pre-trained in the source domain alongside unlabeled data X^t from the target domain to guide the recommendation of annotations. The quantity of target patches chosen for manual annotation is denoted by $N_t^{AL} = \alpha \cdot N_t$, with $N_t^{AL} \ll N_t$ indicating that the actively selected patches are significantly fewer than the total target patches, where α signifies the selection ratio and N_t is the count of all target patches. Concurrently, the labels for these selected patches are symbolized as Y^{Lt} . Our goal is to refine the performance of the model \mathcal{M}^t in the target domain, striving to keep the parameter α as small as possible.

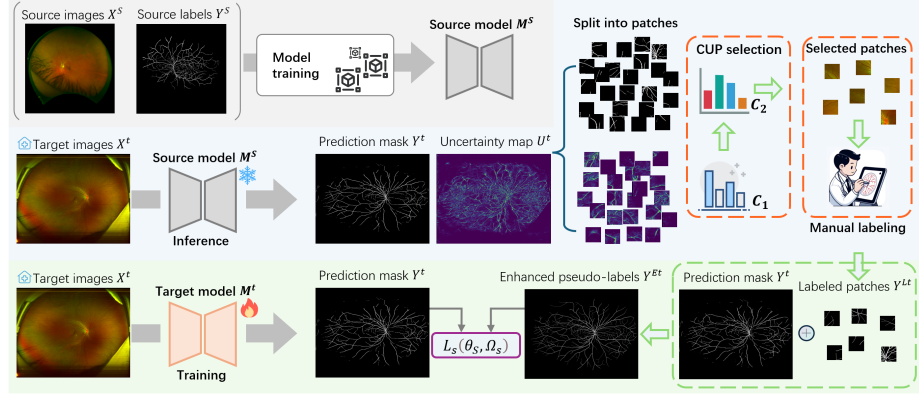


Fig. 2. Overview of the proposed patch-based SFADA framework including the CUP selection strategy. The first row of gray represents the training of source model; the second row of light blue represents the recommendation and annotation of the valuable image patches based on the CUP strategy; and the third row of green represents the target model’s fine-tuning under the supervision of enhanced pseudo-labels.

2.2 Pipeline of Patch-Based SFADA Framework

Considering the high-resolution UWF-SLO images used in our segmentation task, fully annotating the image is extremely expensive and time-consuming. To address this, we propose a method that focuses on selecting and annotating a few valuable image patches and finally integrating them into enhanced pseudo-labels (Y^{Et}) for target model \mathcal{M}^t fine-tuning.

Fig. 2 depicts the structural design of our patch-based SFADA framework. The initial row illustrates the pipeline’s first phase, where we utilize images with their corresponding annotations (X^s, Y^s) from the source domain to train the segmentation model, obtaining the source domain segmentation model \mathcal{M}^s . The second step is depicted in the second row of Fig. 2. Here, we freeze the parameters of \mathcal{M}^s to infer the target domain UWF-SLO image X^t , subsequently deriving its prediction masks Y^t and uncertainty maps U^t (detailed computational methods are explained in Section 2.3). Subsequently, we divide Y^t and U^t into numerous small patches and recommend the most valuable ones to ophthalmologists for annotation, employing our CUP strategy. The third step, illustrated in the third row of Fig. 2, involves merging the actively annotated real patch labels Y^{Lt} with the network’s prediction masks Y^t to produce the enhanced pseudo labels Y^{Et} . The specific operation is to modify the corresponding image patches using Y^{Lt} according to the position index of these patches. Finally, the source domain model \mathcal{M}^s undergoes fine-tuning guided by the enhanced pseudo-label Y^{Et} to develop the target domain model \mathcal{M}^t . This process involves minimizing the loss function $L_s(\theta_s, \Omega_s)$ (a combined loss function of cross-entropy and Dice) in relation to the network parameters Ω_s .

2.3 Cascade Uncertainty-Predominance selection

In the domain adaptation task, the source domain model \mathcal{M}^s has already acquired some fundamental knowledge of vessel segmentation, enabling it to generate a preliminary vessel mask from target domain data X^t . However, the variances between domains result in several regions within the \mathcal{M}^s predicted vessel mask exhibiting high uncertainty. To address this challenge, we design a CUP selection strategy to prioritize patches with higher uncertainty, where \mathcal{M}^s may lack related knowledge. Building on this premise, we further identify patches with substantial vessel prediction by \mathcal{M}^s , denoting regions of predominance. This approach underscores that, despite significant uncertainties in \mathcal{M}^s 's predictions, where model predictions are prone to errors. Certain regions still predict a large number of target vessels, necessitating ground truth annotations for precise model training and refinement.

As illustrated in the CUP selection box of Fig. 2, the CUP strategy comprises two cascades: C_1 for uncertainty and C_2 for predominance. First, we employ the source domain model \mathcal{M}^s to generate the prediction masks Y^t and the uncertainty maps U^t . The procedure for prediction masks is detailed as follows:

$$Prediction_mask = \arg \max_c \left(softmax(\mathcal{M}^s(x^t)) \right), \quad (1)$$

where x^t represents a target image from X^t , and c denotes the predicted category, here refers to the foreground vessels. The process for uncertainty maps is outlined in the following manner:

$$Uncertainty = -1 \times \sum_{c=1}^C softmax(\mathcal{M}^s(x^t)) \log \left(softmax(\mathcal{M}^s(x^t)) \right), \quad (2)$$

where $\sum_{c=1}^C$ represents the sum of c classes, both the foreground vessels and background. This measure reflects the entropy of the model's predictive probability distribution: higher values indicate greater uncertainty (i.e., the model's predictions are spread out across different classes).

Next, as depicted in Fig. 2, the prediction masks Y^t and the uncertainty maps U^t are divided into multiple small patches. For each patch, we calculate the total number of predicted vessel pixels and the aggregate uncertainty values, denoted as $ves_p = \{P_0, \dots, P_t\}$ and $ves_u = \{U_0, \dots, U_t\}$, respectively. Based on these statistical results, our cascade selection strategy is operated as follows:

$$\begin{aligned} \text{step1: } Select_{C1\%} &= \text{Top}(ves_u)[C1\%], \\ \text{step2: } Selected &= \text{Top}(ves_p \text{ in } Select_{C1\%})[C2\%], \end{aligned} \quad (3)$$

where $C1\%$ and $C2\%$ denote the ratio of patches selected based on the highest uncertainty and predicted vessel pixels, respectively.

3 Experiments and Results

3.1 Data Description

We collected 30 UWF-SLO images each from two distinct medical centers, utilizing Optos California and 200Tx cameras (Optos plc, Dunfermline, UK) for capture. The datasets from each center comprised varied categories (refer to Table 1 for details) and were annotated with vessel masks by their respective ophthalmologists. The ophthalmologists utilized Photoshop software for the manual annotation of vessels within the UWF-SLO images. They precisely labeled the vessels across different regions by iteratively fine-tuning the image’s brightness and contrast, adopting a layered approach, and leveraging the software’s outlining tools for accurate delineation. Combined with the currently existing publicly available data PRIME-FP20 [2], we established the first multi-center vessel segmentation of UWF-SLO, with domain shifts potentially attributed to different annotators and different annotation approaches and disease categories.

3.2 Implementation Details and Evaluation Metrics

Implementation Details. For objective evaluation, each dataset is randomly split into three subsets (training, validation, testing) with a ratio of 6:2:2. The model that performs best on the validation set is then chosen for reporting its results on the test set. The PRIME-FP20 [2] dataset is used as the source domain, with centers A and B serving as the target domains. All experiments are carried out on an NVIDIA RTX 3090 GPU with 24 GB memory. The original image size is 3900×3072 and the patch size is set to 260×256 . $C1\%$ and $C2\%$ are set to 10% and 50% respectively, which means that a total of $\alpha = 5\%$ of the patches are selected for annotation. All input images resized to 1024×1024 for uniform training. The SGD optimizer and a batch size of 5 are employed for training. For original training with all labels, models undergo 6000 iterations, while fine-tuning with pseudo labels involves 3000 iterations. An initial learning rate of 0.03 is set, undergoing exponential decay at a rate of 0.9 per iteration. For consistency, comparison methods are re-implemented using the same U-Net backbone [17] and executed under identical conditions.

Evaluation Metrics. Following previous work [15], We employ the Dice score (Dice), Intersection over Union (IoU), Matthews Correlation Coefficient (MCC), and Bookmaker Informedness (BM) as metrics to assess the performance of these models. Higher values indicate superior model performance.

3.3 Experimental Results

This section provides an overview of experimental results across various medical centers in Table 2 and 3, including the lower bound (model without finetuning), upper bound (model finetuned with all labels), and comparisons with other state-of-the-art domain adaptation and active learning methods. Fig. 3 exhibits some visualizations of the segmentation results. By analyzing the data from Table 2

Table 2. Quantitative comparison on Dice and IoU of our method and other state-of-the-art domain adaptation and active learning methods on the MU-VS dataset.

Methods	Dice (mean \pm std, %)			IoU (mean \pm std, %)		
	Center A	Center B	Overall	Center A	Center B	Overall
Lower bound	54.76 \pm 2.73	51.32 \pm 3.25	53.04 \pm 2.99	37.75 \pm 2.56	34.58 \pm 2.91	36.17 \pm 2.74
Upper bound	62.00 \pm 1.32	57.36 \pm 2.99	59.68 \pm 2.16	44.94 \pm 1.38	40.27 \pm 2.96	42.61 \pm 2.17
AdvEnt [23]	56.29 \pm 1.35	51.95 \pm 2.96	54.12 \pm 2.16	39.18 \pm 1.30	35.14 \pm 2.72	37.16 \pm 2.01
DPL [1]	56.50 \pm 2.20	52.58 \pm 2.92	54.54 \pm 2.56	39.40 \pm 2.12	35.72 \pm 2.67	37.56 \pm 2.40
CBMT [20]	57.51 \pm 1.42	52.95 \pm 2.79	55.23 \pm 2.10	40.38 \pm 1.39	36.06 \pm 2.57	38.22 \pm 1.98
CPR [7]	57.79 \pm 2.01	53.28 \pm 2.93	55.54 \pm 2.47	40.66 \pm 1.96	36.36 \pm 2.71	38.51 \pm 2.34
Adversarial [22]	58.79 \pm 1.73	53.32 \pm 2.76	56.06 \pm 2.25	41.66 \pm 1.71	36.40 \pm 2.56	39.03 \pm 2.14
AADA [19]	58.92 \pm 1.46	53.38 \pm 2.84	56.15 \pm 2.15	41.78 \pm 1.46	36.45 \pm 2.66	39.12 \pm 2.06
MHPL [24]	59.32 \pm 1.22	53.58 \pm 2.99	56.45 \pm 2.11	42.18 \pm 1.23	36.66 \pm 2.84	39.42 \pm 2.04
STDR [25]	59.51 \pm 1.54	53.96 \pm 2.80	56.73 \pm 2.17	42.38 \pm 1.56	37.00 \pm 2.67	39.69 \pm 2.11
Ours	60.92 \pm 0.94	54.92 \pm 2.81	57.92 \pm 1.88	43.81 \pm 0.98	37.91 \pm 2.70	40.86 \pm 1.84

Table 3. Quantitative comparison on MCC and BM of our method and other state-of-the-art domain adaptation and active learning methods on the MU-VS dataset.

Methods	MCC (mean \pm std, %)			BM (mean \pm std, %)		
	Center A	Center B	Overall	Center A	Center B	Overall
Lower bound	54.03 \pm 2.54	50.33 \pm 3.25	52.18 \pm 2.90	48.95 \pm 4.91	50.27 \pm 4.74	49.61 \pm 4.83
Upper bound	61.48 \pm 1.29	56.58 \pm 2.80	59.03 \pm 2.05	55.71 \pm 3.98	60.87 \pm 4.23	58.29 \pm 4.10
AdvEnt [23]	55.40 \pm 1.38	50.94 \pm 2.85	53.17 \pm 2.11	51.18 \pm 3.63	52.92 \pm 4.64	52.05 \pm 4.14
DPL [1]	55.69 \pm 2.13	51.60 \pm 2.79	53.65 \pm 2.46	51.07 \pm 4.39	51.39 \pm 4.59	51.23 \pm 4.49
CBMT [20]	56.79 \pm 1.42	51.99 \pm 2.67	54.39 \pm 2.05	51.25 \pm 2.84	52.34 \pm 4.68	51.79 \pm 3.76
CPR [7]	57.21 \pm 1.93	52.35 \pm 2.82	54.78 \pm 2.38	50.88 \pm 3.67	55.33 \pm 4.73	53.11 \pm 4.20
Adversarial [22]	57.99 \pm 1.79	52.39 \pm 2.64	55.19 \pm 2.22	53.27 \pm 3.27	55.60 \pm 4.31	54.44 \pm 3.79
AADA [19]	58.03 \pm 1.54	52.43 \pm 2.70	55.23 \pm 2.12	53.99 \pm 3.00	54.63 \pm 4.88	54.31 \pm 3.94
MHPL [24]	58.38 \pm 1.31	52.72 \pm 2.89	55.55 \pm 2.10	55.00 \pm 2.74	56.91 \pm 5.00	55.96 \pm 3.87
STDR [25]	58.58 \pm 1.64	53.07 \pm 2.67	55.83 \pm 2.16	55.08 \pm 2.97	56.56 \pm 4.78	55.82 \pm 3.88
Ours	59.94 \pm 0.99	54.09 \pm 2.67	57.02 \pm 1.83	57.82 \pm 3.06	57.95 \pm 5.15	57.89 \pm 4.10

and Table 3, it becomes evident that significant performance gaps exist between the lower and upper bounds across various evaluation metrics. For instance, in the case of the BM metric, the overall gap widens from 49.61% to 58.29%.

Domain adaptation analysis. We benchmark our method against the latest state-of-the-art domain adaptation techniques under identical backbone architectures and experimental conditions. This comparison encompasses methods requiring access to source data, such as AdvEnt [23], and source-free approaches like DPL [1], CBMT [20], and CPR [7]. The experimental outcomes indicate that various methods have led to improvements in accuracy. As illustrated in Table 2, the overall Dice scores for these techniques vary from 54.12% to 55.54%, surpassing the lower bound of 53.04%. However, perhaps due to the absence of supervised training with real labels, these improvements are relatively limited, and our method achieved 57.92% in this indicator with few labels (5%).

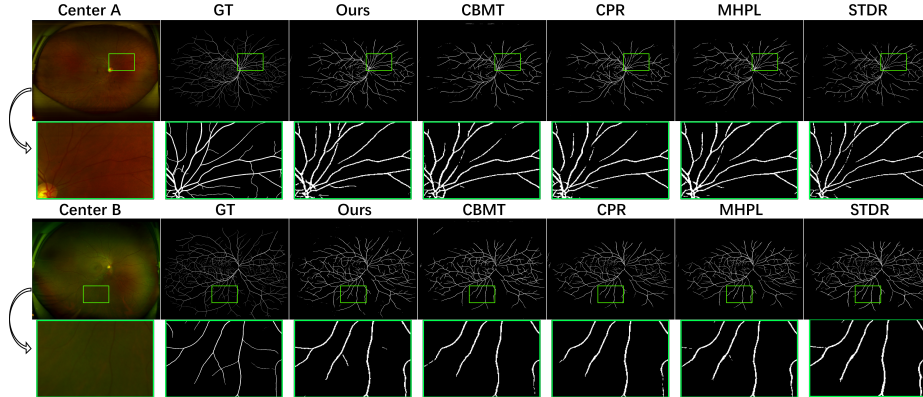
Active learning analysis. Given that SFADA incorporates elements of active learning, we also compare our approach with the recent leading active learning methods, all evaluated under the same experimental setup with 5% labeled data. This comparison includes methods such as Adversarial [22], AADA [19], MHPL [24], and STDR [25], ensuring a comprehensive analysis under uniform condi-

Table 4. Ablation experiments on the MU-VS dataset with Dice and IoU.

Methods	Patch-based Methods			Dice (mean \pm std, %)			IoU (mean \pm std, %)		
	Random	C1	C2	Center A	Center B	Overall	Center A	Center B	Overall
M1	✓	-	-	58.17 \pm 1.89	52.87 \pm 2.78	55.52 \pm 2.34	41.04 \pm 1.86	35.99 \pm 2.56	38.52 \pm 2.21
M2	-	✓	-	59.13 \pm 1.42	53.52 \pm 2.82	56.33 \pm 2.12	41.99 \pm 1.43	36.58 \pm 2.62	39.29 \pm 2.03
Ours	-	✓	✓	60.92 \pm 0.94	54.92 \pm 2.81	57.92 \pm 1.88	43.81 \pm 0.98	37.91 \pm 2.70	40.86 \pm 1.84

Table 5. Ablation experiments on the MU-VS dataset with MCC and BM.

Methods	Patch-based Methods			MCC (mean \pm std, %)			BM (mean \pm std, %)		
	Random	C1	C2	Center A	Center B	Overall	Center A	Center B	Overall
M1	✓	-	-	57.46 \pm 1.88	51.93 \pm 2.66	54.69 \pm 2.27	51.97 \pm 3.51	54.86 \pm 4.80	53.42 \pm 4.15
M2	-	✓	-	58.17 \pm 1.52	52.68 \pm 2.66	55.42 \pm 2.09	54.93 \pm 3.03	56.93 \pm 4.64	55.93 \pm 3.83
Ours	-	✓	✓	59.94 \pm 0.99	54.09 \pm 2.67	57.02 \pm 1.83	57.82 \pm 3.06	57.95 \pm 5.15	57.89 \pm 4.10

**Fig. 3.** Visual comparisons of our method and other leading domain adaptation and active learning methods. Our method can more accurately segment the vessels (please zoom in for more details).

tions. Merging the data from Table 2 and Table 3 reveals that, overall, active learning approaches outperform domain adaptation methods. Notably, our strategy yields the highest scores across all four metrics, underscoring the efficacy of our patch-based approach augmented by the CUP selection strategy.

Ablation Studies. To verify the effectiveness of our method, we conduct corresponding ablation experiments (the results are shown in Table 4 and Table 5), including three configurations as follows: (1) M1: randomly select patches combined with our patch-based framework. (2) M2: based on $C1$ uncertainty, the top 5% patches are selected for annotation and then integrated into the overall framework. (3) Ours: performing cascade selection, first select $C1$ uncertainty and then $C2$ predominance, named the CUP strategy. The experimental results show that M2 outperforms M1 overall, while Ours achieves greater performance gains compared to M2, e.g., the overall Dice from 55.52% to 56.33%, and finally to 57.92% in Table 4, which proves the effectiveness of our CUP strategy.

4 Conclusion

In this paper, we explore the task of vessel segmentation of UWF-SLO images across different centers. Considering the high cost of labeling high-resolution UWF-SLO images, we propose a patch-based SFADA approach to significantly save labeling resources while boosting segmentation performance. We also devise a CUP strategy to cascade the selection of valuable patches with high uncertainty and dominance for annotation. Meanwhile, we construct the first public multi-center UWF-SLO vessel segmentation (MU-VS) dataset to facilitate related research. Experimental results demonstrate that our method achieves optimal results compared to other domain adaptation and active learning methods. In the future, we plan to extend our method to other similar high-resolution medical image segmentation tasks.

Acknowledgements This work is supported by the Guangzhou-HKUST(GZ) Joint Funding Program (No. 2023A03J0671), the Guangzhou Municipal Science and Technology Project (Grant No. 2023A03J0671), NSFC Key Project U23A20391, NSFC General Project 62072452, and the InnoHK funding launched by Innovation and Technology Commission, Hong Kong SAR.

Disclosure of Interests The authors declare that they have no competing interests.

References

1. Chen, C., Liu, Q., Jin, Y., Dou, Q., Heng, P.A.: Source-free domain adaptive fundus image segmentation with denoised pseudo-labeling. In: MICCAI. pp. 225–235. Springer (2021)
2. Ding, L., Kuriyan, A.E., Ramchandran, R.S., Wykoff, C.C., Sharma, G.: Weakly-supervised vessel detection in ultra-widefield fundus photography via iterative multi-modal registration and learning. *IEEE Transactions on Medical Imaging* **40**(10), 2748–2758 (2020)
3. Fraz, M.M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A.R., Owen, C.G., Barman, S.A.: Blood vessel segmentation methodologies in retinal images—a survey. *Computer methods and programs in biomedicine* **108**(1), 407–433 (2012)
4. Galdran, A., Costa, P., Bria, A., Araújo, T., Mendonça, A.M., Campilho, A.: A no-reference quality metric for retinal vessel tree segmentation. In: International conference on medical image computing and computer-assisted intervention. pp. 82–90. Springer (2018)
5. Guan, H., Liu, M.: Domain adaptation for medical image analysis: a survey. *IEEE Transactions on Biomedical Engineering* **69**(3), 1173–1185 (2021)
6. Hoover, A., Kouznetsova, V., Goldbaum, M.: Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical imaging* **19**(3), 203–210 (2000)
7. Huai, Z., Ding, X., Li, Y., Li, X.: Context-aware pseudo-label refinement for source-free domain adaptive fundus image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 618–628. Springer (2023)

8. Kumari, S., Singh, P.: Deep learning for unsupervised domain adaptation in medical imaging: Recent advancements and future perspectives. *Computers in Biology and Medicine* p. 107912 (2023)
9. Li, X., Hao, H., Fu, H., Zhang, D., Chen, D., Qiao, Y., Liu, J., Zhao, Y., Zhang, J.: Privileged modality guided network for retinal vessel segmentation in ultra-wide-field images. In: *International Workshop on Ophthalmic Medical Image Analysis*. pp. 82–91. Springer (2023)
10. Luo, Z., Luo, X., Gao, Z., Wang, G.: An uncertainty-guided tiered self-training framework for active source-free domain adaptation in prostate segmentation (2024), <https://arxiv.org/abs/2407.02893>
11. Mookiah, M.R.K., Hogg, S., MacGillivray, T.J., Prathiba, V., Pradeepa, R., Mohan, V., Anjana, R.M., Doney, A.S., Palmer, C.N., Trucco, E.: A review of machine learning methods for retinal blood vessel segmentation and artery/vein classification. *Medical Image Analysis* **68**, 101905 (2021)
12. Nagiel, A., Lalane, R.A., Sadda, S.R., Schwartz, S.D.: Ultra-widefield fundus imaging: a review of clinical applications and future trends. *Retina* **36**(4), 660–678 (2016)
13. Niu, Z., Ouyang, S., Xie, S., Chen, Y.w., Lin, L.: A survey on domain generalization for medical image analysis. *arXiv preprint arXiv:2402.05035* (2024)
14. Pellegrini, E., Robertson, G., Trucco, E., MacGillivray, T.J., Lupascu, C., van Hemert, J., Williams, M.C., Newby, D.E., van Beek, E.J., Houston, G.: Blood vessel segmentation and width estimation in ultra-wide field scanning laser ophthalmoscopy. *Biomedical optics express* **5**(12), 4329–4337 (2014)
15. Qiu, Z., Hu, Y., Chen, X., Zeng, D., Hu, Q., Liu, J.: Rethinking dual-stream super-resolution semantic learning in medical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023)
16. Rogers, S., McIntosh, R.L., Cheung, N., Lim, L., Wang, J.J., Mitchell, P., Kowalski, J.W., Nguyen, H., Wong, T.Y., Consortium, I.E.D., et al.: The prevalence of retinal vein occlusion: pooled data from population studies from the united states, europe, asia, and australia. *Ophthalmology* **117**(2), 313–319 (2010)
17. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *MICCAI*. pp. 234–241 (2015)
18. Staal, J., Abramoff, M.D., Niemeijer, M., Viergever, M.A., Van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging* **23**(4), 501–509 (2004)
19. Su, J.C., Tsai, Y.H., Sohn, K., Liu, B., Maji, S., Chandraker, M.: Active adversarial domain adaptation. In: *WACV*. pp. 739–748 (2020)
20. Tang, L., Li, K., He, C., Zhang, Y., Li, X.: Source-free domain adaptive fundus image segmentation with class-balanced mean teacher. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 684–694. Springer (2023)
21. Tang, Q.Q., Yang, X.G., Wang, H.Q., Wu, D.W., Zhang, M.X.: Applications of deep learning for detecting ophthalmic diseases with ultrawide-field fundus images. *International Journal of Ophthalmology* **17**(1), 188 (2024)
22. Tsai, Y.H., Hung, W.C., Schuler, S., Sohn, K., Yang, M.H., Chandraker, M.: Learning to adapt structured output space for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 7472–7481 (2018)
23. Vu, T.H., Jain, H., Bucher, M., Cord, M., Pérez, P.: Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. In: *CVPR*. pp. 2517–2526 (2019)

24. Wang, F., Han, Z., Zhang, Z., He, R., Yin, Y.: Mhpl: Minimum happy points learning for active source free domain adaptation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 20008–20018 (2023)
25. Wang, H., Chen, J., Zhang, S., He, Y., Xu, J., Wu, M., He, J., Liao, W., Luo, X.: Dual-reference source-free active domain adaptation for nasopharyngeal carcinoma tumor segmentation across multiple hospitals. *IEEE Transactions on Medical Imaging* (2024)
26. Wang, H., et al.: Advancing delineation of gross tumor volume based on magnetic resonance imaging by performing source-free domain adaptation in nasopharyngeal carcinoma. In: International Workshop on Computational Mathematics Modeling in Cancer Analysis. pp. 71–80. Springer (2023)
27. Xing, Z., Wan, L., Fu, H., Yang, G., Zhu, L.: Diff-unet: A diffusion embedded network for volumetric segmentation. *arXiv preprint arXiv:2303.10326* (2023)
28. Xing, Z., Ye, T., Yang, Y., Liu, G., Zhu, L.: Segmamba: Long-range sequential modeling mamba for 3d medical image segmentation. *arXiv preprint arXiv:2401.13560* (2024)
29. Xu, R., Liu, T., Ye, X., Lin, L., Chen, Y.W.: Boosting connectivity in retinal vessel segmentation via a recursive semantics-guided network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 786–795. Springer (2020)
30. Zhang, N., Lu, J., et al.: Source-free unsupervised domain adaptation: Current research and future directions. *Neurocomputing* p. 126921 (2023)
31. Zhao, H., Dong, W., Yu, R., Zhao, Z., Bo, D., Xu, Y.: Morestyle: Relax low-frequency constraint of fourier-based image reconstruction in generalizable medical image segmentation. *arXiv preprint arXiv:2403.11689* (2024)