Test-Time Training with Local Contrast-Preserving Copy-Pasted Image for Domain Generalization in Retinal Vessel Segmentation

Yuliang Gu, Zhichao Sun, Zelong Liu, and Yongchao Xu✉[0000−−0002−−7253−−3151]

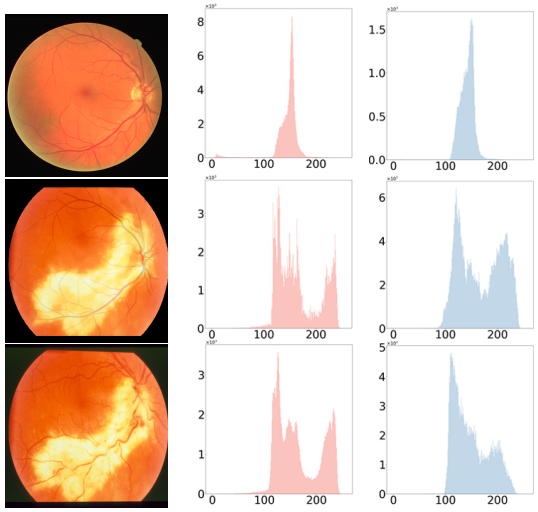
National Engineering Research Center for Multimedia Software, Institute of Artificial Intelligence, School of Computer Science, Medical Artificial Intelligence Research Institute of Renmin Hospital, Wuhan University, Wuhan, China
yongchao.xu@whu.edu.cn

Abstract. Accurate segmentation of retinal vessels is an important task. Deep learning-based approaches have achieved impressive segmentation performance on images with the same distribution as the training images. However, the performance significantly drops when there is a substantial disparity between the distributions of the training and testing data, which limits the practical applicability of these methods in real-world scenarios. In this paper, we propose a novel test-time training (TTT) strategy that employs a local contrast-preserving copy-paste (L2CP) method to generate synthetic images in the target domain style. Specifically, leveraging the thin nature of retinal vessel structures, we apply a simple morphological closing to remove these structures from the test image. This process yields a vessel-free image that retains the target domain's style, which we then employ as the background component for the synthetic image. To realistically integrate retinal vessels from source domain images into the background component, our L2CP method pastes the local contrast map of the vessels, rather than their grayscale values, onto the background component. This approach effectively mitigates the issue of significant disparities in grayscale distribution between the foreground and background across the source and target domains. Extensive TTT experiments on retinal vessel segmentation tasks demonstrate that the proposed L2CP consistently improves the model's generalization ability in retinal structure segmentation. The code of our implementation is available at https://github.com/GuGuLL123/L2CP.

Keywords: Test Time Training · Domain Generalization · Vessel Segmentation.

1 Introduction

Accurate segmentation of blood vessels is very important in many applications. For example, an essential prerequisite for computational hemodynamics and retinal fundus disease screening is the accurate segmentation of blood vessels [3].



(a) Image (b) Background (c) Foreground

Fig. 1: Image, Background distribution, and foreground distribution of source domain (first row), synthetic domain (second row), and target domain (third row) images. Distribution difference between synthetic domain and target domain is significantly reduced compared to source-target domain discrepancy. The histogram is obtained from the grayscale image.

Retinal vessel segmentation poses specific challenges compared to general object segmentation [16]. Firstly, these structures are often thin and long, requiring specialized segmentation techniques for accurate extraction. Secondly, low contrast (in particular for these thin vessels) between vessels and the background complicates segmentation. Lastly, diverse appearances in color, texture, and background illumination present a significant challenge, requiring segmentation algorithms to address and correct these changes.

Classical methods [12,11] focus on designing filters for extracting features related to vessel structures, often requiring careful parameter adjustments. Recent deep learning methods [16,14,17,7,9,2,22,10,13,18,29,27,24] achieve remarkable results and can be categorized into: 1) Designing networks for the elongated structure of vessels [14,13]; 2) Incorporating prior knowledge of vessels [16,22]; 3) Formulating loss functions specific to vessel shape and topology [7,18,29]. Despite impressive intra-dataset results, the challenge of domain generalization often arises in real-world applications. For instance, notable distribution discrepancies may arise between testing and training images due to ocular lesions, see Fig. 1. Many methods that perform well within the domain encounter challenges in effectively generalizing to out-of-distribution (OOD) scenarios [16,30]. Test-time training (TTT) methods [1,21,23] aim at bridging the gap between the training and testing datasets during the testing phase. TENT [21] reduces

entropy-based loss to update the batch normalization layers for test samples batch-wise. CoTTA [23] enables long-term adaptation for network parameters to addresses the challenges of non-stationary environments in test-time domain adaptation. However, these TTT methods are not specifically designed for the thin nature of retinal vessels, leading to poor performance in cross-domain retinal vessel segmentation.

In this paper, we generate local contrast-preserving copy-pasted image as a bridge between the source domain and target domain to fine-tune model during testing, addressing the domain generalization problem for retinal vessel segmentation. Leveraging the prior knowledge that the retinal vessels are thin, we utilize morphological closing to eliminate vessels from the target domain images. This process results in vessel-free images preserving the target domain's style, which we subsequently use as the background component for the synthetic image. Subsequently, we propose a novel L2CP method to copy-paste the local contrast maps in source domain images onto vessel-free background, yielding realistic synthetic images that contain vessels of source domain images and background preserving the target domain's style. As show in Fig. 1, the L2CP method effectively alleviates the problem of the large gap in both background and foreground grayscale distribution between source and target domain. Fine-tuning the model using these synthetic images during testing is effective in enhancing the model's capability to handle out-of-distribution data. We conduct TTT experiments using these synthetic images on three retinal blood vessel segmentation dataset [19,6,4], resulting in significant and consistent performance enhancements for the domain generalization task.

The main contributions of this paper are threefold. 1) We propose a simple and novel strategy to realistically integrate vessels in source domain into the target domain background component. The proposed L2CP effectively alleviates the problem of the large gap in both background and foreground grayscale distribution between source and target domain. 2) We leverage synthetic image as a bridge to fine-tune the model during testing. Furthermore, we conduct extensive experiments in the field of generalizable retinal blood vessel segmentation using multiple classical networks and datasets. The proposed L2CP consistently enhances the model's generalization capability in vessel structure segmentation.

2 Method

2.1 Test-time training with synthetic image

Problem setting. Given a model f_{θ_0} pretrained by source domain data $\mathcal{D}^S = \{(I^S, G^S)\}$ (I^S represents image and G^S represents corresponding ground truth), our goal is to enhance the performance of the current model for target domain data $\mathcal{D}^T = \{I^T\}$ during inference time in an online manner. According to the one-pass protocol [20], the testing samples are sequentially processed, and a model update occurs after each testing sample is fed. At time step t, target image I_t^T is given as input, we fine-tune the model's partial parameters $\theta_{t-1} \rightarrow \theta_t$ according to I_t^T. The final prediction of I_t^T is predicted by f_{θ_t}.

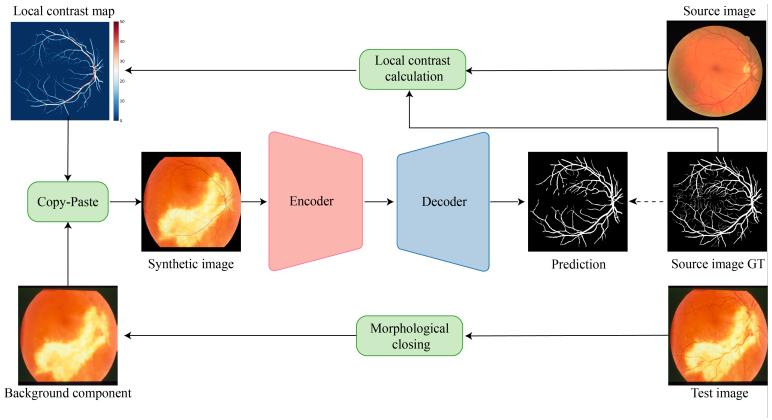


Fig. 2: The pipeline of the proposed test-time training with local contrast-preserving copy-paste (L2CP) strategy. The morphological closing is applied to remove vessels from the test image. The proposed L2CP ensures vessels originating from the source domain are realistically pasted into the test image.

Overview of L2CP. The pipeline of the proposed framework is depicted in Fig 2. The local contrast map C^S is calculated using one source domain data (I^S, G^S). When testing target domain image I_t^T at time step t, we employ mathematical morphology to remove vessels in an unsupervised manner, obtaining the background component $\bar{I}_t^T$. L2CP Copy-pastes C^S into $\bar{I}_t^T$, generating the synthetic image I_t^{L2CP} that contains vessels corresponding to G^S and background preserving the target domain's style. Subsequently, we use G^S as the ground truth of I_t^{L2CP} to fine-tune $\theta_{t-1} \rightarrow \theta_t$. Finally, we use f_{θ_t} to get the segmentation of I_t^T. The proposed method requires only a minimal amount of source domain information (only one image and its corresponding annotation of retinal blood vessel dataset), which adheres to the sTTT protocol [20] that allowing access to a light-weight information from the source domain.

2.2 Generation vessel-free image of target domain style.

The inherent thinness of vessel structures results in their relatively low pixel occupancy in images. Leveraging such inherent characteristics, we apply morphological closing operation [8] to remove vessel structures in a given test image. Since vessels are often relatively darker than the background [16], we employ morphological closing operations for image processing. Using kernel $\mathcal{O} = \{ (u, v) \mid u \in [-k, k],\ v \in [-k, k]\}$ with size k to perform morphological closing operation on an target image I_t can be denoted as $\bar{I}^T = I^T \bullet \mathcal{O} = (I^T \oplus \mathcal{O}) \ominus \mathcal{O}$, where $I^T \oplus \mathcal{O}(x, y) = \max_{(u,v)\in\mathcal{O}}\{I^T(x-u, y-v)\}$ represents dilation operation and $I^T \ominus \mathcal{O}(x, y) = \min_{(u,v)\in\mathcal{O}}\{I^T(x-u, y-v)\}$ represents erosion operation.

2.3 Local contrast-preserving copy-paste

Given the significant disparities in data distribution across domains, directly pasting the gray values of vessels from the source domain into $\bar{I}^T$ might result in the situation that these structures' gray values surpass those of the local background. This circumstance could undermine the effectiveness of the underlying assumption that vessels inherently possess characteristics of being darker than the context, potentially affecting the neural network's fine-tuning process. To ensure consistency in the local contrast between vessel structures and the surrounding background in the synthetic image and the source domain image, we propose L2CP method. We define $\mathcal{V} = \{(x, y) \mid G^S(x, y) = 1\}$ as the set of coordinates representing the vessels. For the position $(x, y) \in \mathcal{V}$ of vessels in the source domain, we calculate the local contrast map by:

$$C^S(x, y) = I^S(x, y) - \frac{1}{|\mathcal{O}_{\mathcal{V}}(x, y)|} \sum_{(u,v)\in\mathcal{O}_{\mathcal{V}}(x,y)} I^S(x - u, y - v), \tag{1}$$

where $\mathcal{O}_{\mathcal{V}}(x, y) = \{(u, v)|(u, v) \in \mathcal{O}, (x - u, y - v) \notin \mathcal{V}\}$ means local background pixels around (x, y) and $|\cdot|$ denotes the cardinality. When $(x, y) \notin \mathcal{V}$, we define $C^S(x, y) = 0$.

For the position $(x, y) \in \mathcal{V}$ of vessels in the background removal image $\bar{I}^T$, we calculate the mean gray value in local background $\mathcal{O}_{\mathcal{V}}(x, y)$:

$$B^T(x, y) = \frac{1}{|\mathcal{O}_{\mathcal{V}}(x, y)|} \sum_{(u,v)\in\mathcal{O}_{\mathcal{V}}(x,y)} \bar{I}^T(x - u, y - v). \tag{2}$$

When $(x, y) \notin \mathcal{V}$, we define $B^T(x, y) = \bar{I}^T(x, y)$.

The synthetic image I^{L2CP} using L2CP is defined as:

$$I^{L2CP}(x, y) = \begin{cases} B^T(x, y) + C^S(x, y) & (x, y) \in \mathcal{V} \\ \bar{I}^T(x, y) & (x, y) \notin \mathcal{V} \end{cases} \tag{3}$$

Using Eq. (1) - Eq. (3) , for $(x, y) \in \mathcal{V}$, we have the following property:

$$I^{L2CP}(x, y) - \frac{1}{|O_{\mathcal{V}}|} \sum_{(u,v)\in O_T} I^{L2CP}(x - u, y - v) = C^S(x, y). \tag{4}$$

This property guarantees that the synthetic image maintains a same local contrast map with the source domain image.

3 Experiments

3.1 Datasets and evaluation metrics

We evaluate the proposed method on three public datasets to demonstrate the generalizability of the proposed method. **DRIVE** [19] dataset consists of 20

Table 1: Comparison results of the proposed L2CP and TTT methods under different baselines for cross-dataset evaluation on retinal vessel segmentation.

Cross-dataset	CHASEDB1 ⇓ DRIVE		CHASEDB1 ⇓ STARE		DRIVE ⇓ STARE		DRIVE ⇓ CHASEDB1		STARE ⇓ DRIVE		STARE ⇓ CHASEDB1		Average
Methods	AUC	F1	AUC	F1	AUC	F1	AUC	F1	AUC	F1	AUC	F1	F1
UNet [15]	0.957	0.729	0.966	0.753	0.942	0.699	0.967	0.758	0.938	0.735	0.945	0.711	0.730
+Tent [21]	0.951	0.735	0.961	0.748	0.951	0.713	0.967	0.758	0.905	0.712	0.943	0.711	0.729
+Cotta [23]	0.957	0.744	0.964	0.754	0.944	0.716	0.968	0.761	0.929	0.718	0.949	0.715	0.734
+Dplot [26]	0.949	0.741	0.964	0.760	0.954	0.710	0.967	0.761	0.913	0.714	0.940	0.712	0.733
+L2CP	**0.961**	**0.763**	**0.970**	**0.761**	**0.959**	**0.729**	**0.972**	**0.773**	**0.944**	**0.745**	**0.950**	**0.720**	**0.749**
LadderNet [31]	0.961	0.739	0.970	0.747	**0.958**	0.719	0.964	0.751	0.964	0.748	0.902	0.693	0.732
+Tent [21]	0.955	0.730	0.944	0.715	0.940	0.711	0.955	0.734	0.957	0.719	0.887	0.660	0.711
+Cotta [23]	0.960	0.719	0.967	0.722	0.944	0.716	0.968	**0.761**	0.963	0.716	0.895	0.667	0.716
+Dplot [26]	0.952	0.706	0.960	0.707	0.909	0.688	0.946	0.733	0.958	0.718	0.893	0.679	0.705
+L2CP	**0.965**	**0.769**	**0.973**	**0.759**	0.956	**0.725**	**0.968**	0.760	**0.965**	**0.771**	**0.932**	**0.720**	**0.751**
LIOT [16]	0.966	0.686	0.979	0.759	0.978	0.773	0.966	0.726	0.958	0.702	0.939	0.593	0.706
+Tent [21]	**0.967**	0.709	**0.981**	0.773	0.977	0.771	0.965	0.716	**0.959**	0.713	0.941	0.618	0.717
+Cotta [23]	0.967	0.705	0.979	0.764	0.978	0.771	**0.968**	**0.744**	0.959	0.714	0.940	0.639	0.722
+Dplot [26]	0.967	0.713	0.980	0.772	0.976	0.774	0.967	0.736	0.960	0.717	0.945	0.633	0.724
+L2CP	0.966	**0.735**	0.980	**0.784**	**0.980**	**0.775**	0.966	0.728	0.957	**0.745**	**0.942**	**0.672**	**0.740**

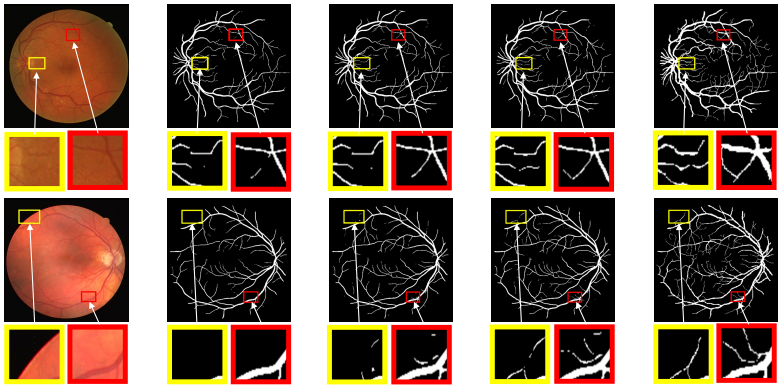
Table 2: Quantitative results of L2CP generalizing from different sampling (superscript) of varying numbers (subscript) of images.

F1			AUC		
Exp_1^1	Exp_1^2	Exp_1^3	Exp_1^1	Exp_1^2	Exp_1^3
0.7712	0.7693	0.7707	0.9650	0.9639	0.9644
Exp_2^1	Exp_2^2	Exp_2^3	Exp_2^1	Exp_2^2	Exp_2^3
0.7718	0.7712	0.7708	0.9648	0.9650	0.9646
Exp_3^1	Exp_3^2	Exp_3^3	Exp_3^1	Exp_3^2	Exp_3^3
0.7710	0.7722	0.7712	0.9644	0.9649	0.9645

train images and 20 test images with a resolution of 565×584 pixels. **STARE** [6] consists of 20 700×605 retinal vessel images, which are divided into 10 training and 10 test images. **CHASEDB1** [4] consists of 28, 999×960 color retinal images, split into 20 training images and 8 test images. We evaluate L2CP on retinal vessel segmentation following the two evaluation metrics [16]: area under the receiver operating characteristics curve (AUC), and F1-score.

3.2 Implementation details

To ensure the experimental process closely simulates the real-world application scenario, all experiments strictly adhere to the one-pass protocol [20]. This involves a single training iteration where inference is immediately conducted on each sample. We use only one source domain image throughout all tests to derive local contrast maps of the blood vessels. The experiments include one-step fine-tuning process, using a batch size of 1. Throughout the adaptation phase,



(a) Image (b) UNet (c) Dplot (d) L2CP (e) GT

Fig. 3: Some qualitative segmentation results of L2CP and some other methods under cross-dataset evaluation on retinal images

we deploy the Adam optimizer with a learning rate set to 0.0001. The kernel size k is set to 6 in all experiments.

3.3 Results

Tab. 1 depicts the quantitative results. In each cross-dataset experiment, we include three classical vessel segmentation methods, UNet [15], LadderNet [31], and LIOT [16] as baseline models. These methods are augmented with four TTT methods, Tent [21], Cotta [23],Dplot [26] and L2CP for experimental comparison. The table clearly demonstrates the consistent superiority of L2CP over others in nearly all experimental settings. Utilizing three baseline models, the proposed L2CP achieves average improvements of 1.9% (UNet), 1.9% (LadderNet) and 3.4% (LIOT) under six experimental settings. Specifically, the proposed method achieves 3.4% increase in the F1 score compared to UNet for the generation from CHASEDB1 dataset [4] to the DRIVE dataset [19]. When evaluating from STARE [6] to the CHASEDB1 dataset [4], L2CP shows a remarkable 7.9% enhancement in the F1 score compared with LIOT. Some qualitative results are shown in Fig. 3, where L2CP achieves accurate segmentation results.

3.4 Ablation studies

Impact of using different source domain images. We employ the notation Exp_i^j to represent experiments involving the selection of distinct source domain images, where i denotes the number of source domain images chosen for the current experiment, and j represents the $j-th$ random selection. Tab. 2 shows the quantitative results using different source domain images. From the experimental results, it is evident that the usage of different samples and quantities of

Table 3: Evaluation of the proposed L2CP using different iteration rounds during test-time training. 0 means the baseline result without test-time training.

Iteration	0	1	2	3	4	5	6	7	8
F1	0.729	0.771	0.776	0.774	0.772	0.771	0.772	0.770	0.769

Table 4: F1 score of generalizing results when using different method to generate synthetic image.

Method	Baseline	Source	Mix-up [28]	Copy-Paste [5]	Fourier [25]	L2CP
F1	0.748	0.754	0.724	0.731	0.750	0.771

source domain images does not impact L2CP 's generalization performance in retinal blood vessel segmentation tasks. L2CP necessitates only one source domain image. All ablation studies are conducted from STARE dataset to DRIVE dataset under LadderNet baseline.

Impact of number of iteration rounds. We conduct ablation experiments on using different iteration rounds during testing. The experimental results in Tab. 3 indicate that the outcome after one single iteration closely approximates the results obtained with multiple iterations, affirming the efficiency of our method. All remaining experiments in this paper are consistently carried out with one single iteration.

Impact of using different synthetic methods. To validate the efficacy of our proposed method, we conduct ablation experiments using different techniques to generate synthetic images. The results are presented in Tab. 4. *Source* denotes results obtained by fine-tuning directly with source domain images. *Mixup* [28] and *Copy-paste* [5] are two useful data augmentation technologies to generate images. *Fourier* indicates results attained by training after replacing the low-frequency components of the amplitude spectrum of target domain images with those of source domain images [25]. As shown in Tab. 4, the comparison methods show no enhancement of the baseline results, demonstrating the effectiveness of employing L2CP-synthesized images for TTT.

4 Conclusion

In this paper, we propose a novel L2CP method in the Test-Time Training scenario to generate synthetic images for domain generalization in retinal vessel segmentation. Specifically, utilizing prior knowledge of the thin nature of vessel structures, we leverage the morphological closing to effectively remove vessels from the test images, forming the background component in the synthetic image. Additionally, the proposed L2CP method alleviates the problem of the large gap in both background and foreground grayscale distribution between the source domain and the target domain, ensuring that vessels originating from the source domain are realistically pasted onto the background component. L2CP

necessitates the retention of a minimal number of source domain images. During testing, a single step of network fine-tuning is adequate, rendering it an efficient and highly effective approach for test-time training. The limitation of the current framework is the requirement of one annotated source domain images. In the future, we would like to use fractal methods [17] to synthesize vessel structures without relying on source-domain data within TTT framework. We will also extend L2CP to other retinal conditions (*e.g.*, diabetic retinopathy).

Acknowledgments. This work was supported in part by the NSFC 62176186 and 62222112, the NSF of Hubei Province of China (2024AFB245), the Special Fund for Central Guidance on Local Science and Technology Development from the Sichuan Provincial Department of Science and Technology (2024ZYD0285), the Key Research and Development Program of Dazhou Science and Technology Bureau (24ZDYF0005).

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

References

1. Chen, Z., Pan, Y., Ye, Y., Lu, M., Xia, Y.: Each test image deserves a specific prompt: Continual test-time adaptation for 2d medical image segmentation. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. pp. 11184–11193 (2024)
2. Deng, Z., Xu, S., Zhang, J., Zhang, J., Wang, D.J., Yan, L., Shi, Y.: Shape-aware 3d small vessel segmentation with local contrast guided attention. In: Proc. of Intl. Conf. on Medical Image Computing and Computer Assisted Intervention. pp. 354–363 (2023)
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. Comput. Methods Programs Biomed. **108**(1), 407–433 (2012)
4. Fraz, M.M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A.R., Owen, C.G., Barman, S.A.: An ensemble classification-based approach applied to retinal blood vessel segmentation. IEEE Transactions on Biomedical Engineering **59**(9), 2538–2548 (2012)
5. Ghiasi, G., Cui, Y., Srinivas, A., Qian, R., Lin, T.Y., Cubuk, E.D., Le, Q.V., Zoph, B.: Simple copy-paste is a strong data augmentation method for instance segmentation. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. pp. 2918–2928 (2021)
6. Hoover, A., Kouznetsova, V., Goldbaum, M.: Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Trans. on Medical Imaging **19**(3), 203–210 (2000)
7. Hu, X., Li, F., Samaras, D., Chen, C.: Topology-preserving deep image segmentation. Proc. of Advances in Neural Information Processing Systems **32** (2019)
8. Jähne, B.: Digital image processing. Springer Science & Business Media (2005)
9. Kreitner, L., Paetzold, J.C., Rauch, N., Chen, C., Hagag, A.M., Fayed, A.E., Sivaprasad, S., Rausch, S., Weichsel, J., Menze, B.H., et al.: Synthetic optical coherence tomography angiographs for detailed retinal vessel segmentation without human annotations. IEEE Trans. on Medical Imaging (2024)

10. Liu, Y., Yu, B., Chen, T., Gu, Y., Du, B., Xu, Y., Cheng, J.: Progressive retinal image registration via global and local deformable transformations. In: IEEE International Conference on Bioinformatics and Biomedicine. pp. 2183–2190 (2024)
11. Memari, N., Ramli, A.R., Saripan, M.I.B., Mashohor, S., Moghbel, M.: Retinal blood vessel segmentation by using matched filtering and fuzzy c-means clustering with integrated level set method for diabetic retinopathy assessment. Journal of Medical and Biological Engineering **39**, 713–731 (2019)
12. Merveille, O., Talbot, H., Najman, L., Passat, N.: Curvilinear structure analysis by ranking the orientation responses of path operators. IEEE Trans. on Pattern Anal. and Mach. Intell. **40**(2), 304–317 (2017)
13. Mou, L., Zhao, Y., Fu, H., Liu, Y., Cheng, J., Zheng, Y., Su, P., Yang, J., Chen, L., Frangi, A.F., et al.: CS2-Net: Deep learning segmentation of curvilinear structures in medical imaging. Medical image analysis **67**, 101874 (2021)
14. Qi, Y., He, Y., Qi, X., Zhang, Y., Yang, G.: Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. pp. 6070–6079 (2023)
15. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: Proc. of Intl. Conf. on Medical Image Computing and Computer Assisted Intervention. pp. 234–241 (2015)
16. Shi, T., Boutry, N., Xu, Y., Géraud, T.: Local intensity order transformation for robust curvilinear object segmentation. IEEE Trans. on Image Processing **31**, 2557–2569 (2022)
17. Shi, T., Ding, X., Zhang, L., Yang, X.: FreeCOS: Self-supervised learning from fractals and unlabeled images for curvilinear object segmentation. In: Proc. of IEEE Intl. Conf. on Computer Vision. pp. 876–886 (2023)
18. Shit, S., Paetzold, J.C., Sekuboyina, A., Ezhov, I., Unger, A., Zhylka, A., Pluim, J.P., Bauer, U., Menze, B.H.: clDice-a novel topology-preserving loss function for tubular structure segmentation. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. pp. 16560–16569 (2021)
19. Staal, J., Abràmoff, M.D., Niemeijer, M., Viergever, M.A., Van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. IEEE Trans. on Medical Imaging **23**(4), 501–509 (2004)
20. Su, Y., Xu, X., Jia, K.: Revisiting realistic test-time training: Sequential inference and adaptation by anchored clustering. Proc. of Advances in Neural Information Processing Systems **35**, 17543–17555 (2022)
21. Wang, D., Shelhamer, E., Liu, S., Olshausen, B., Darrell, T.: Tent: Fully test-time adaptation by entropy minimization. In: Proc. of International Conference on Learning Representations (2021)
22. Wang, D., Zhang, Z., Zhao, Z., Liu, Y., Chen, Y., Wang, L.: PointScatter: Point set representation for tubular structure extraction. In: Proc. of European Conf. on Computer Vision. pp. 366–383 (2022)
23. Wang, Q., Fink, O., Van Gool, L., Dai, D.: Continual test-time domain adaptation. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. pp. 7201–7211 (2022)
24. Xu, J., Dong, A., Yang, Y., Jin, S., Zeng, J., Xu, Z., Jiang, W., Zhang, L., Dong, J., Wang, B.: VSNet: Vessel structure-aware network for hepatic and portal vein segmentation. Medical Image Analysis (2025)
25. Xu, Q., Zhang, R., Zhang, Y., Wang, Y., Tian, Q.: A fourier-based framework for domain generalization. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. pp. 14383–14392 (2021)

26. Yu, Y., Shin, S., Back, S., Ko, M., Noh, S., Lee, K.: Domain-specific block selection and paired-view pseudo-labeling for online test-time adaptation. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. pp. 22723–22732 (2024)
27. Zhang, B., Zhang, Z., Liu, S., Faghihroohi, S., Schunkert, H., Navab, N.: XA-Sim2Real: Adaptive representation learning for vessel segmentation in x-ray angiography. In: Proc. of Intl. Conf. on Medical Image Computing and Computer Assisted Intervention (2024)
28. Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: MIXUP: Beyond empirical risk minimization. In: Proc. of International Conference on Learning Representations (2018)
29. Zhang, J., Qiao, Y., Sarabi, M.S., Khansari, M.M., Gahm, J.K., Kashani, A.H., Shi, Y.: 3D shape modeling and analysis of retinal microvasculature in oct-angiography images. IEEE Trans. on Medical Imaging **39**(5), 1335–1346 (2019)
30. Zhou, K., Liu, Z., Qiao, Y., Xiang, T., Loy, C.C.: Domain generalization: A survey. IEEE Trans. on Pattern Anal. and Mach. Intell. (2022)
31. Zhuang, J.: LadderNet: Multi-path networks based on u-net for medical image segmentation. arXiv preprint arXiv:1810.07810 (2018)