A Multi-Branch Framework for Cross-Domain Vessel Segmentation via the Few-Shot Paradigm

Zihang Huang, Tianyu Zhao, Liang Zhang, and Xin Yang (✉)

School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, China
xinyang2014@hust.edu.cn

Abstract. In recent years, deep learning-based vessel segmentation methods have made significant progress. However, the diversity of image modalities and the high-cost of acquiring sufficient annotated data constrain the performance of existing approaches. Given that the primary objective of segmenting various types of vessels is to extract high-frequency tubular structures, leveraging existing annotated datasets for training and fast generalizing to novel vessel segmentation tasks is an ideal solution to the above challenges, which can be achieved by the few-shot segmentation (FSS) paradigm. Unfortunately, the significant differences in texture and thickness among different types of vessels leave unsolved challenges. To address this issue, we propose a novel framework that incorporates FSS into cross-domain vessel segmentation. In particular, we construct high-frequency auxiliary modalities to guide the model in focusing on high-frequency features, which are highly correlated with vessel regions, thereby bridging the texture gap between images of various vessel types. Furthermore, we design a Dual-Modal Feature Extraction and Fusion (DM-FEF) module to extract modality-specific features. Finally, addressing the thickness variations between different vessels, we designed a Multi-Branch Feature Extractor (MBFE) module to capture the diverse characteristics of vessels with different thickness, enabling the model to perceive the thickness differences between distinct vessels. Experimental results on six public datasets demonstrate the effectiveness of our method. Source code: https://github.com/ZiH-Huang/FSS_Cross.

Keywords: Medical image · Vessel segmentation · Few-shot learning.

1 Introduction

Accurate vessel segmentation plays a crucial role in the diagnosis and treatment of related diseases [21, 20]. However, the diversity of image modalities, limited data availability, and the high-cost of annotation constrain the performance of existing methods. In the medical imaging field, to address these challenges, few-shot segmentation (FSS) [19] becomes a research focus, enabling pre-trained models to segment unseen classes using limited prompts without additional training. This paradigm achieves a better balance between accuracy and efficiency compared to unsupervised[23, 9, 10] or weakly supervised methods[14], which is

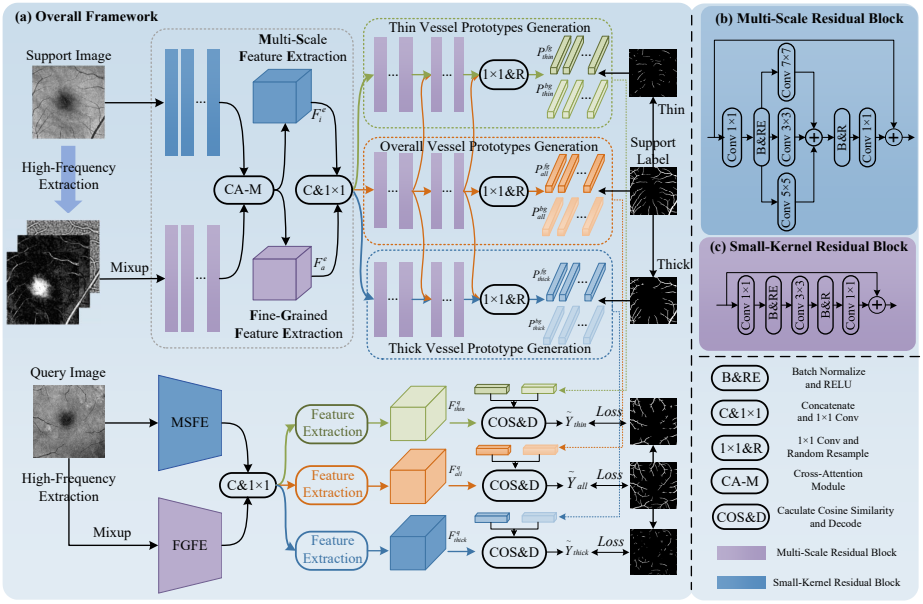

Fig. 1. The overall framework of the proposed method. The encoders for the support image and the query image share the same parameters.

well-suited for vessel segmentation, as it allows training on existing datasets (e.g., public color fundus datasets [26, 12, 8]), and when novel data types are introduced (e.g., private coronary artery data). The trained model only requires a few annotated samples as support to segment target vessels, eliminating the need for retraining. Unfortunately, the significant differences in imaging modalities and target characteristics across various vessel datasets (i.e., the cross-domain problem) present substantial challenges for applying FSS to cross-domain vessel segmentation tasks, and this issue remains largely unexplored.

To solve the above problems, we propose a novel framework to incorporate FSS into cross-domain vessel segmentation, and with a focus on mitigating performance degradation caused by the **differences in texture and vessel thickness** between distinct types of vessels [23]. Firstly, based on the commonality among different vessel segmentation tasks in the extraction of high-frequency tubular structures, we propose a High-Frequency Auxiliary Modality (HFAM) construction method. By combining traditional salience detection [2, 28] and Laplacian transformation, HFAM effectively highlights the high-frequency details, guides the model to focus on curvilinear structural characteristics, and reduces the **texture gap** in cross-domain vessel segmentation. This auxiliary modality image contains more high-frequency information compared to the original image, while the original image retains more low-frequency information related to the background. Based on this, we designed a Dual-Modal Feature Extraction and Fusion (DM-FEF) module. In DM-FEF, two branches with

different structures are employed to capture unique information from varying modalities. Additionally, we introduce the cross-attention mechanism between the two branches to enhance feature interaction, allowing the model to better integrate complementary information. Finally, traditional few-shot segmentation approaches [6, 31] tend to extract one or multiple representative prototypes in a single feature space to represent the segmentation target, making it challenging to capture the feature variations between vessels of different thicknesses. Accordingly, we designed a Multi-Branch Feature Extractor (MBFE) module. Through utilizing independent branches to capture the features of vessels with varying thicknesses, the model gains the ability to perceive vessel thickness. This approach helps mitigate the negative impact of **thickness variations** in different vessel types.

The contributions of our research are as follows: (1) We propose the HFAM method, which effectively highlights high-frequency details and reduces the texture gap in cross-domain vessel segmentation. (2) We design the DM-FEF module, enabling the encoder to capture unique information from both HFAM and the original images. (3) We introduce the MBFE module, allowing the model to perceive vessel thickness and mitigate the negative impact of thickness variations across different vessel types. (4) To the best of our knowledge, we are the first to incorporate FSS into the cross-domain vessel segmentation task. Comparison results with 10 state-of-the-art (SOTA) methods across 6 public datasets demonstrate the superiority of the proposed approach.

2 Method

In this paper, the dataset is divided into the training set D_{train} and the testing set D_{test}, which belong to different domains. Specifically, the training dataset D_{train} consists of image-label pairs $\{(X, Y)\}_{train}$. In each training iteration, a random subset $S_{train} = \{(X_i, Y_i)\}_{i=1}^{S}$, is selected as the support set, where S is the number of support samples. S_{train} guides the model to segment samples in the query set Q_{train}, where $S_{train} \cap Q_{train} = \emptyset$. During testing, a random subset of S image-label pairs is selected from D_{test} to guide the segmentation of the remaining samples. In this study, the number of support samples S is set to 1. The framework is shown in Fig. 1, with the main modules described as follows.

2.1 High-frequency auxiliary modality

Given the importance of high-frequency information in capturing structural and textural features, we propose to create high-frequency auxiliary modality (HFAM) for effectively highlighting vessel-related regions. Firstly, we employ traditional low-cost saliency detection methods, including FT [1] and HC [4] to generate saliency images X_f and X_h that highlight high-frequency regions and suppress low-frequency background noise. Meanwhile, to enhance edge features, we apply Laplacian transformation to emphasize object contours and critical structures, and obtain the edge-enhanced image X_l. Inspired by Mixup [32], to

fully leverage the complementary characteristics of X_f, X_h and X_l, and construct diverse training samples, we fuse them according to the following formula:

$$X_{HFAM} = \lambda_1 \times X_f + \lambda_2 \times X_h + \lambda_3 \times X_l \tag{1}$$

here λ_1, λ_2, and λ_3 are random weights generated from a Dirichlet distribution, the probability density function is expressed as:

$$f(\lambda_1, \lambda_2, \lambda_3; \boldsymbol{\alpha}) = \frac{\Gamma\left(\sum_{i=1}^{3} \alpha_i\right)}{\prod_{i=1}^{3} \Gamma(\alpha_i)} \prod_{i=1}^{3} \lambda_i^{\alpha_i - 1} \tag{2}$$

here the concentration parameter is $\alpha = [1.5, 1.5, 1.5]$, $\Gamma(\cdot)$ is the gamma function. X_{HFAM} enhances the saliency regions and edge features of X, amplifying the high-frequency information and critical structures relevant to the target.

2.2 Dual-modal feature extraction and fusion

Auxiliary modality images capture rich high-frequency information, while original images include both foreground details and low-frequency background. Accordingly, at the early stage of feature extraction, we design a Dual-Modal Feature Extraction and Fusion (DM-FEF) module with three main components: 1) a multi-scale residual feature extraction branch to capture local details and global contextual information from original images. 2) a fine-grained feature extraction branch to extract high-frequency information from auxiliary modality images. 3) cross-attention enhancement modules to strengthen feature alignment. The module is based on the *conv*2_x layer of ResNet-50 [11].

Multi-Scale Feature Extraction. The residual blocks are modified by incorporating multi-scale parallel convolutions of different kernel sizes (3×3, 5×5, and 7×7) for low- and high-frequency feature extraction. Small kernels capture local details, while large kernels capture broader context.

Fine-Grained Feature Extraction. Given the nature of auxiliary modality images, we select small convolution kernels to focus on details like edge [29], preserving high-frequency features. 3×3 convolution is chosen as the small convolution block. The structure remains consistent with *conv*2_x in ResNet-50.

Cross-Attention Enhancement Module. To facilitate effective feature fusion, we introduce cross-attention for dynamic interaction between the branches. Taking the multi-scale feature extracting branch as an example, given the original image feature F_i, we initially transform the input sequences into three distinct representations K_i, Q_i and V_i, the attention matrix A_i can be computed as:

$$A_i = \text{softmax}(Q_i \cdot K_a^\top), \;\; V_i^* \;\; = V_a \cdot A_i^\top, \;\; F_i^e = \gamma \cdot V_i^* + F_i \tag{3}$$

where K_a, V_a is the distinct representations obtained from fine-grained feature extraction branch, and γ is a learnable parameter initialized to zero. Then we fuse the enhanced features of the two branches by 1×1 convolution:

$$F_f = RELU(BN(Conv_{1\times1}(F_i^e \oplus F_a^e))) \tag{4}$$

where $\oplus$ represents the concatenation operation, F_a^e is the features obtained from fine-grained feature extraction branch and BN is batch normalization.

2.3 Multi-Branch Feature Extraction

Thickness variations between vessel types hinder cross-domain segmentation. For example, XCAD [18] vessels are typically thicker than OCTA [16] vessels, causing the model to prioritize thick vessel features and struggle with segmenting thin vessels in OCTA, and vice versa. To address this, we follow [30] to separate the entire vessels into thick and thin vessels, and introducing two independent branches B_{thick} and B_{thin} to learn their features in distinct spaces, enabling the model to perceive thickness differences. Meanwhile, we introduced an additional branch B_{all} to provide global cues for B_{thick} and B_{thin}, assisting them in accurately distinguishing and localizing the overall vascular structure. The structure of each branch is consistent with layers $conv3\_x$, and $conv4\_x$ in ResNet-50. The features of support and query samples obtained from B_x are denoted as $F_x^s, F_x^q \in \mathbb{R}^{C\times H\times W}$, respectively, where $x \in \{thin, thick, all\}$. H, W are the height and width of the obtained features, and C denotes the channels the features ($C = 256$ in this paper). For the support features F_x^s, we follow [6] to randomly sample N ($N = 400$ in this paper) multiple representative background and foreground prototypes P_x^{bg}, $P_x^{fg} \in N \times C$. We then compute the cosine similarity between each spatial location (h, w) in F_x^q and the n-th background and foreground prototypes:

$$\begin{aligned} Q_x^{bg}(n,h,w) &= \frac{F_x^q(:,h,w)\cdot P_x^{bg}(n,:)^\top}{\|F_x^q(:,h,w)\|_2 \cdot \|P_x^{bg}(n,:)\|_2} \\ Q_x^{fg}(n,h,w) &= \frac{F_x^q(:,h,w)\cdot P_x^{fg}(n,:)^\top}{\|F_x^q(:,h,w)\|_2 \cdot \|P_x^{fg}(n,:)\|_2} \end{aligned} \tag{5}$$

here $Q_x^{bg}, Q_x^{fg} \in \mathbb{R}^{N\times H\times W}$. Then, we adopt two independent decoders, each composed of a few convolutional layers, to fuse the similarity results:

$$\tilde{Q}_x^{bg} = Decoder(Q_x^{bg}, \theta_x^{bg}),\ \ \tilde{Q}_x^{fg} = Decoder(Q_x^{fg}, \theta_x^{fg}) \tag{6}$$

where $\tilde{Q}_x^{bg}, \tilde{Q}_x^{fg} \in \mathbb{R}^{H\times W}$, and θ_x^{bg} and θ_x^{fg} are the parameters of the decoders. The final prediction logits $\tilde{Y}_x$ can be calculated by a softmax function:

$$\tilde{Y}_x = softmax(\tilde{Q}_x^{bg} \oplus \tilde{Q}_x^{fg}) \tag{7}$$

where $\tilde{Y}_x \in \mathbb{R}^{2\times H\times W}$ and $\oplus$ represents the concatenation operation.

2.4 Loss Function

In this paper, we employ Dice loss and Cross-Entropy loss to measure the distance between predictions and the ground-truth. For thin vessel segmentation, we compute L_{thin}^{ce} and L_{thin}^{dice}, excluding thick vessel regions. Similarly, for thick vessel segmentation, thin vessel regions are ignored. The overall loss is:

$$L_{total} = \alpha \times (L_{thin}^{ce} + L_{thin}^{dice}) + (L_{thick}^{ce} + L_{thick}^{dice)} + (L_{all}^{ce} + L_{all}^{dice}) + L_{mse} \tag{8}$$

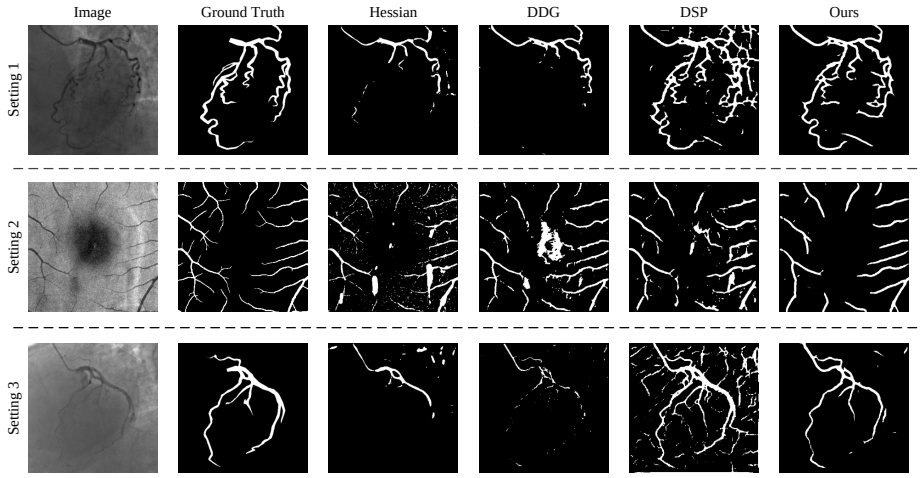

Fig. 2. Visualization of segmentation results of the proposed method and SOTA.

here we set α= 2.0 to make the model focus more on thin vessels, which more difficult to be segmented, and L_{mse} is the Mean Squared Error (MSE) loss, which ensures distinct predictions for thick and thin vessels in the foreground. During the model inference stage, we combine the outputs of the thick vessel prediction branch B_{thick} and the thin vessel prediction branch B_{thin} to obtain the final vessel segmentation result.

3 Experiments

3.1 Dataset and evaluation metrics

We adopt six public datasets, including: (1) Three color fundus datasets (Color Fundus): DRIVE (40 samples) [26], STARE (20 samples) [12], and CHASE_DB1 (28 samples) [8]. (2) Two X-ray angiography datasets: XCAD (126 samples) [18] and XCA (134 samples) [3]. (3) Optical Coherence Tomography Angiography ataset [16], which contains two subsets: OCTA_6M (300 samples) and OCTA_3M (200 samples). We evaluate our method in three cross-domain scenes: (1) Training on OCTA_3M and testing on XCAD. (2) Training on XCAD and XCA, and testing on OCTA_3M and OCTA_6M. (3) Training on three color fundus datasets and testing on XCAD. We adopt common evaluation metrics in the experiments: Dice Coefficient (Dice), Centerline Dice Coefficient (clDice) [25], Sensitivity (Sn), Specificity (Sp), and Accuracy (Acc).

3.2 Implementation details

In experiments, each sample is resized to 400 × 400 pixels. We adopt Adam optimizer (weight decay = 0.00001, batch size = 1) and the initial learning rate

Table 1. The results of comparison experiments in three cross domain scenes.

Method		OCTA_3M → XCAD					XCAD/XCA → OCTA_3M/6M					Color Fundus → XCAD				
		clDice	Sn	Sp	Acc	Dice	clDice	Sn	Sp	Acc	Dice	clDice	Sn	Sp	Acc	Dice
Traditional	OOF	34.22	27.73	96.93	92.94	29.60	46.64	55.16	93.81	90.56	47.72	34.22	27.73	96.93	92.94	29.60
	Hessian	48.50	45.02	96.94	93.95	45.06	45.41	45.35	96.49	92.28	47.90	48.50	45.02	96.94	93.95	45.06
Generalization	LIOT	50.49	59.21	94.34	92.26	45.96	39.51	34.79	97.71	92.49	42.81	46.78	62.13	94.07	92.15	46.68
	Affinity	38.27	28.86	98.31	94.29	34.21	41.69	36.76	**97.73**	92.38	42.71	46.14	47.17	**96.95**	94.03	46.08
	DDG	45.59	36.03	**98.32**	94.71	41.45	47.89	46.24	96.10	91.94	47.78	48.63	52.62	96.72	94.13	49.81
	VFT	-	-	-	-	-	-	-	-	-	-	46.57	63.55	95.84	93.93	46.57
Few-shot	CAT	44.68	45.24	96.17	93.17	41.75	48.10	48.08	93.40	89.26	45.09	42.92	56.44	95.37	93.03	47.22
	GMRD	47.82	50.89	96.33	93.67	47.12	49.98	51.80	93.08	89.32	46.87	48.29	54.41	96.29	93.84	48.29
	PAMI	46.34	54.86	94.39	92.12	42.92	53.71	53.45	94.47	91.79	46.71	53.70	70.87	92.17	90.88	47.04
	DSP	47.95	60.28	94.31	92.30	51.39	51.39	55.77	92.27	89.27	46.41	45.84	**81.62**	91.36	90.75	45.84
	Ours	**61.45**	**65.91**	97.04	**95.18**	**60.31**	**58.28**	**59.11**	94.87	**92.52**	**50.91**	**57.32**	68.00	95.81	**94.14**	**56.18**

is 0.001, which is doubled in the first two epochs and then decays exponentially. The training process lasts for 40 epochs. All experiments are performed on one NVIDIA 3090 GPU with 24 GB memory, the CUDA version is 11.7 and the Pytorch version is 13.0. In testing, the support sample is randomly selected, and the results of each setting are average over three tests. The vessels are divided into thick and thin following [24] and [30]. When evaluate the final results, since the outcome from B_{all} is relatively rough compared to B_{thin} and B_{thick}, the results from B_{thin} and B_{thick} are combined for evaluation.

3.3 Comparison with the state-of-the-arts methods

We compare our method with ten state-of-the-art methods (SOTA), including the traditional methods: Hessian [7] and OOF [15], the generalization methods: LIOT [22], Affinity [24], VFT [13], DDG [5], and few-shot methods: CAT [17], GMRD [6], PAMI [33], DSP [27]. The results are shown in Table 1 and the visualization results are shown in Fig. 2. The experiments show that our method surpasses all compared methods across all three settings. In setting 1, it achieves a Dice of 60.31% and clDice of 61.45%, exceeding SOTAs by 8.92% and 10.96%, respectively. For setting 2, the method attains a Dice of 50.91% and clDice of 58.28%, outperforming SOTAs by 4.57% and 3.01%. In setting 3, it achieves a Dice of 56.18% and clDice of 57.32%, surpassing SOTAs by 6.37% and 3.62%.

3.4 Ablation study

We conduct several ablation studies to evaluate each component:

High-frequency auxiliary modality. We conduct experiments in **setting** 1 and 2 to verify the proposed HFAM. DM-FEF and MBFE modules are adopted. The experiments include: (1) Only using original images for training (Original). (2) Adopting saliency detection (Salience)/Laplacian transformed images (Laplacian) to assist feature extraction. (3) The proposed Mixup-based method (Mix). As shown in Table 2, Mix significantly improves model performance.

Dual-modal feature extraction and fusion module. We conduct experiments to evaluate the DM-FEF module. The HFAM and MBFE module is adopted. The experiments include: (1) We adopt small kernels with sharing

Table 2. The ablation study of three components: HFAM, DM-FEF and MBFE.

Method		OCTA_3M → XCAD					XCAD/XCA → OCTA_3M/6M				
		clDICE	Sn	Sp	Acc	Dice	clDICE	Sn	Sp	Acc	Dice
HFAM	Original	56.75	61.43	95.95	93.91	52.66	53.21	54.29	93.61	90.39	47.53
	Salience	56.26	66.16	95.46	93.68	53.61	54.25	55.37	93.09	89.65	49.26
	Laplacian	58.09	**66.48**	95.88	94.13	55.59	55.42	54.91	93.87	90.40	50.56
	Mix	**61.45**	65.91	**97.04**	**95.18**	**60.31**	**58.28**	**59.11**	**94.87**	**92.52**	**50.91**
DM-FEF	Sharing	57.69	60.79	96.38	94.24	54.21	55.69	56.78	93.50	92.49	49.13
	Individual	54.88	65.90	95.72	93.95	55.62	56.24	57.08	94.41	91.51	48.47
	Multi	61.03	63.73	96.80	94.83	57.73	57.16	58.12	94.74	92.32	50.23
	DM-FEF	**61.45**	**65.91**	**97.04**	**95.18**	**60.31**	**58.28**	**59.11**	**94.87**	**92.52**	**50.91**
MBFE	Single	52.42	54.41	96.29	93.84	49.47	51.57	52.26	93.86	90.08	48.85
	Two	59.32	62.57	96.67	94.63	56.37	54.68	55.93	94.79	92.33	49.14
	MBFE	**61.45**	**65.91**	**97.04**	**95.18**	**60.31**	**58.28**	**59.11**	**94.87**	**92.52**	**50.91**

parameters (Sharing)/ individual parameters (Individual). (2) We adopt multi-scale kernels in two branches with individual parameters (Multi). (3) The proposed DM-FEF module. As shown in Table 2, DM-FEF achieves the best results.

Multi-branch feature extraction. The experiments include: (1) Adopting a single branch to segment all vessels (Single). (2) Adopting two branches to segment thin and thick vessels without B_{all} (Two). (3) The proposed MBFE module. As shown in Table 2, through enabling the model to distinguish thin and thick vessels, clDice improves by 6.90% and 3.11% and Dice increases 6.90% and 0.29%, respectively. B_{all} enhances the performance, with clDice further increasing 2.13% and 3.94% and Dice improving by 3.94% and 1.77%, respectively.

Overall framework. We exhibit the ablation study of the overall framework in OCTA_3M → XCAD, shown in Table 3. Considering the proposed method is based on multi-descriptors, we use results of GMRD [6] as the baseline results. Compared to baseline, the performance of the proposed method shows significant improvement with clDice increasing by 13.63% and Dice improving by 13.19%.

Table 3. The ablation study of the overall framework.

Baseline	HFAM	DMFE-FM	MBFE	clDice	Sn	Sp	Acc	Dice
√				47.82	50.89	96.33	93.67	47.12
√	√			51.53	53.61	96.11	93.58	48.04
√	√	√		52.42	54.41	96.29	93.84	49.47
√			√	56.75	61.43	95.95	93.91	52.66
√	√	√	√	**61.45**	**65.91**	**97.04**	**95.18**	**60.31**

4 Conclusion

In this work, we propose a novel multi-branch framework for few-shot cross-domain vessel segmentation. The framework introduces High-Frequency Auxil-

iary Modality (HFAM) to highlight vessel-related regions, eliminating the texture gap in cross domain scenes. Meanwhile, a Dual-Modal Feature Extraction and Fusion (DM-FEF) module is proposed to capture the modality-specific features. To mitigate the negative impact of thickness variations in different vessel types, we proposed a Multi-Branch Feature Extractor (MBFE) to equip the model with the ability to perceive vessel thickness. Experiments on six public datasets demonstrate the superiority of our method.

Acknowledgement. This work was supported in part by the National Natural Science Foundation of China under Grant 62472184, and in part by the Fundamental Research Funds for the Central Universities.

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

References

1. Achanta, R., Hemami, S., Estrada, F., Susstrunk, S.: Frequency-tuned salient region detection. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 1597–1604. IEEE (2009)
2. Borji, A., Cheng, M.M., Hou, Q., Jiang, H., Li, J.: Salient object detection: A survey. Computational visual media **5**, 117–150 (2019)
3. Cervantes-Sanchez, F., Cruz-Aceves, I., Hernandez-Aguirre, A., Hernandez-Gonzalez, M.A., Solorio-Meza, S.E.: Automatic segmentation of coronary arteries in x-ray angiograms using multiscale analysis and artificial neural networks. Applied Sciences **9**(24), 5507 (2019)
4. Cheng, M.M., Mitra, N.J., Huang, X., Torr, P.H., Hu, S.M.: Global contrast based salient region detection. IEEE transactions on pattern analysis and machine intelligence **37**(3), 569–582 (2014)
5. Cheng, Z., Liu, M., Yan, C., Wang, S.: Dynamic domain generalization for medical image segmentation. Neural Networks **184**, 107073 (2025)
6. Cheng, Z., Wang, S., Xin, T., Zhou, T., Zhang, H., Shao, L.: Few-shot medical image segmentation via generating multiple representative descriptors. IEEE Transactions on Medical Imaging (2024)
7. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: Medical Image Computing and Computer-Assisted Intervention—MICCAI'98: First International Conference Cambridge, MA, USA, October 11–13, 1998 Proceedings 1. pp. 130–137. Springer (1998)
8. 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)
9. Gao, S., Fu, Y., Liu, K., Gao, W., Xu, H., Wu, J., Han, Y.: Collaborative knowledge amalgamation: Preserving discriminability and transferability in unsupervised learning. Information Sciences **669**, 120564 (2024)
10. Gao, S., Fu, Y., Liu, K., Han, Y.: Contrastive knowledge amalgamation for unsupervised image classification. In: International Conference on Artificial Neural Networks. pp. 192–204. Springer (2023)

11. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
12. 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)
13. Hu, D., Li, H., Liu, H., Oguz, I.: Domain generalization for retinal vessel segmentation via hessian-based vector field. Medical Image Analysis **95**, 103164 (2024)
14. Huang, Z., Wang, Z., Zhao, T., Ding, X., Yang, X.: Toward high-quality pseudo masks from noisy or weak annotations for robust medical image segmentation. Neural Networks **181**, 106850 (2025)
15. Law, M.W., Chung, A.C.: Three dimensional curvilinear structure detection using optimally oriented flux. In: Computer Vision–ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part IV 10. pp. 368–382. Springer (2008)
16. Li, M., Chen, Y., Yuan, S., Chen, Q.: Octa-500. IEEE Transactions on Medical Imaging **39**(9), 2806–2818 (2019). https://doi.org/10.1109/TMI.2020.2992244, https://doi.org/10.1109/TMI.2020.2992244
17. Lin, Y., Chen, Y., Cheng, K.T., Chen, H.: Few shot medical image segmentation with cross attention transformer. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 233–243. Springer (2023)
18. Ma, Y., Hua, Y., Deng, H., Song, T., Wang, H., Xue, Z., Cao, H., Ma, R., Guan, H.: Self-supervised vessel segmentation via adversarial learning. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 7536–7545 (2021)
19. Pachetti, E., Colantonio, S.: A systematic review of few-shot learning in medical imaging. Artificial intelligence in medicine p. 102949 (2024)
20. Pu, J., Gezer, N.S., Ren, S., Alpaydin, A.O., Avci, E.R., Risbano, M.G., Rivera-Lebron, B., Chan, S.Y.W., Leader, J.K.: Automated detection and segmentation of pulmonary embolisms on computed tomography pulmonary angiography (ctpa) using deep learning but without manual outlining. Medical image analysis **89**, 102882 (2023)
21. Qin, Q., Chen, Y.: A review of retinal vessel segmentation for fundus image analysis. Engineering Applications of Artificial Intelligence **128**, 107454 (2024)
22. Shi, T., Boutry, N., Xu, Y., Géraud, T.: Local intensity order transformation for robust curvilinear object segmentation. IEEE Transactions on Image Processing **31**, 2557–2569 (2022)
23. Shi, T., Ding, X., Zhang, L., Yang, X.: Freecos: self-supervised learning from fractals and unlabeled images for curvilinear object segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 876–886 (2023)
24. Shi, T., Ding, X., Zhou, W., Pan, F., Yan, Z., Bai, X., Yang, X.: Affinity feature strengthening for accurate, complete and robust vessel segmentation. IEEE Journal of Biomedical and Health Informatics **27**(8), 4006–4017 (2023)
25. 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: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 16560–16569 (2021)
26. Staal, J., Abràmoff, 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)

27. Tang, S., Yan, S., Qi, X., Gao, J., Ye, M., Zhang, J., Zhu, X.: Few-shot medical image segmentation with high-fidelity prototypes. Medical Image Analysis **100**, 103412 (2025)
28. Wang, W., Lai, Q., Fu, H., Shen, J., Ling, H., Yang, R.: Salient object detection in the deep learning era: An in-depth survey. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(6), 3239–3259 (2021)
29. Wang, X., Xu, G., Jia, H., Yang, X.: Selective-stereo: Adaptive frequency information selection for stereo matching. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 19701–19710 (2024)
30. Yan, Z., Yang, X., Cheng, K.T.: A three-stage deep learning model for accurate retinal vessel segmentation. IEEE journal of Biomedical and Health Informatics **23**(4), 1427–1436 (2018)
31. Zhang, B., Xiao, J., Qin, T.: Self-guided and cross-guided learning for few-shot segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8312–8321 (2021)
32. Zhang, H.: mixup: Beyond empirical risk minimization. arXiv preprint arXiv:1710.09412 (2017)
33. Zhu, Y., Wang, S., Xin, T., Zhang, Z., Zhang, H.: Partition-a-medical-image: Extracting multiple representative sub-regions for few-shot medical image segmentation. IEEE Transactions on Instrumentation and Measurement (2024)