GLCP: Global-to-Local Connectivity Preservation for Tubular Structure Segmentation

Feixiang Zhou[1], Zhuangzhi Gao[1], He Zhao[1], Jianyang Xie[1], Yanda Meng[2], Yitian Zhao[3], Gregory Y.H. Lip[4], and Yalin Zheng[1]([✉])

[1] Department of Eye and Vision Sciences, University of Liverpool, Liverpool, UK
yalin.zheng@liverpool.ac.uk
[2] Department of Computer Science, University of Exeter, Exeter, UK
[3] Ningbo Institute of Materials Technology and Engineering, CAS, Ningbo, China
[4] Department of Cardiovascular and Metabolic Medicine, University of Liverpool, Liverpool, UK

Abstract. Accurate segmentation of tubular structures, such as vascular networks, plays a critical role in various medical domains. A remaining significant challenge in this task is structural fragmentation, which can adversely impact downstream applications. Existing methods primarily focus on designing various loss functions to constrain global topological structures. However, they often overlook local discontinuity regions, leading to suboptimal segmentation results. To overcome this limitation, we propose a novel Global-to-Local Connectivity Preservation (GLCP) framework that can simultaneously perceive global and local structural characteristics of tubular networks. Specifically, we propose an Interactive Multi-head Segmentation (IMS) module to jointly learn global segmentation, skeleton maps, and local discontinuity maps, respectively. This enables our model to explicitly target local discontinuity regions while maintaining global topological integrity. In addition, we design a lightweight Dual-Attention-based Refinement (DAR) module to further improve segmentation quality by refining the resulting segmentation maps. Extensive experiments on both 2D and 3D datasets demonstrate that our GLCP achieves superior accuracy and continuity in tubular structure segmentation compared to several state-of-the-art approaches. The source codes will be available at https://github.com/FeixiangZhou/GLCP.

Keywords: Tubular structure segmentation · Vascular segmentation · Connectivity preservation.

1 Introduction

The precise segmentation of thin tubular structures, such as blood vessels, is a fundamental step for many downstream tasks, including disease diagnosis [16], surgical planning [1] and computational biology [8]. However, segmenting tubular structures is challenging due to their elongated, thin shapes, branching patterns, as well as imaging imperfections such as poor contrast.
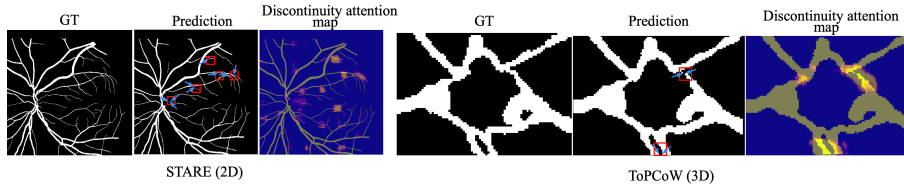
Fig. 1. Illustration of motivation. In most existing methods, structural fragmentation (red boxes) commonly occurs in predicted masks, manifesting as redundant endpoints (indicated by blue arrows). We propose to directly focus on these regions during training by adaptively learning discontinuity maps, allowing the model to better capture local structural characteristics for connectivity enhancement. For clear visualization, we use the max intensity projection maps of the 3D ground truth (GT) and prediction.

Existing deep learning segmentation methods can be grouped as model-based, feature-based, and loss-based segmentation approaches. Model-based methods [3,15,18,26] focus on tailoring network architectures to align with the unique characteristics of tubular structures. However, due to the inherent sparsity of tubular structures in images, these methods often struggle to accurately capture and delineate them. Feature-based methods [13,7,27,17] aim to capture the geometric and topological characteristics of tubular structures by incorporating additional feature representations into the model, but their performance and efficiency may be compromised by redundant feature representations. Loss-based methods [20,11,23,19,14,28] develop different loss functions to strengthen the topological connectivity of tubular structures, such as topological constraints based on persistent homology [2,6] or skeleton-based formulations [20,19]. Despite these advances, current methods tend to focus on global topological constraints while neglecting the fine-grained, local characteristics of discontinuity-prone areas. In [12], a topology violation map identifies topological errors in local regions, but its effectiveness in 3D segmentation is uncertain. Additionally, requiring an extra network to refine predictions increases its overall computational complexity. Recent work [7] has designed a multi-task paradigm to enhance global connectivity and boundary consistency. However, the interactions between these tasks are limited, and their attention to discontinuities remains insufficient. Therefore, how to identify and correct structural fragmentation, especially in regions prone to discontinuities, remains an open and challenging problem.

In this paper, we proposed a GLCP framework tailored for both 2D and 3D tasks. Instead of applying topological constraints to the predictions, we proposed Interactive Multi-head Segmentation (IMS) that jointly learns global segmentation, skeleton maps, and local discontinuity maps by a multi-head end-to-end architecture. Specifically, in addition to the segmentation task, we first designed a novel discontinuity prediction task, where a discontinuity head is introduced to predict discontinuity-prone local regions. Intuitively, the structural fragmen-

tation in vascular segmentation results often manifests as redundant endpoints, as shown in Fig. 1. This motivates us to detect these endpoints from the initial predictions (i.e., segmentation maps) to identify potential discontinuity regions, as shown in Fig. 2, thereby guiding the model to better focus on local structures by adaptively learning these discontinuities (i.e., discontinuity maps). We also constructed a skeleton prediction task by adding a skeleton head to learn global skeletal representation (i.e., skeleton maps) of foreground objects. In particular, we propose a self-supervised consistency loss to boost the interactions between the main task and skeleton prediction task. Equipped with this multi-head architecture and a shared backbone, our method can better capture global and local structural characteristics of tubular structures, thus improving both segmentation accuracy and topological continuity.

Apart from the IMS, we also introduced a lightweight Dual-Attention-based Refinement (DAR) module to refine the segmentation results based on the initial segmentation, discontinuity and skeleton maps. More specifically, the probability maps derived from the skeleton and discontinuity maps are utilized as global and local attention maps, guiding the model to refine critical regions, thereby improving the overall segmentation quality.

Our contributions are as follows: **1)** We propose an IMS strategy to enhance the model's capability to simultaneously perceive global and local structural characteristics of tubular networks, leading to better connectivity. **2)** We design a DAR module to refine segmentation results by integrating global and local attention mechanisms based on skeleton and discontinuity maps. **3)** Extensive experiments on both 2D and 3D datasets demonstrate that our approach outperforms existing methods in terms of both segmentation accuracy and topological continuity.

2 Methodology

An overview of the proposed method is illustrated in Fig. 2. Three interactive heads (i.e., H_g, H_d and H_s) are designed to jointly learn segmentation, discontinuity and skeleton maps (i.e., $\hat{F}_g$, $\hat{F}_d$ and $\hat{F}_s$), allowing our model to pay attention to both global and local structures (Sec. 2.1). Besides, discontinuity and skeleton attention maps extracted from the corresponding predictions are utilized to further refine the segmentation results (Sec. 2.2).

2.1 Interactive Multi-head Segmentation (IMS)

As aforementioned, existing methods lack the ability to perceive local discontinuity regions. To address this limitation, we propose a novel auxiliary task to predict potential discontinuity regions, which enhances the model's awareness of the discontinuities and the robustness of segmentation results.

Specifically, a new discontinuity head H_d is integrated into the backbone to produce the discontinuity maps $\hat{F}_d$. However, to enable adaptive learning, the corresponding GT should be carefully constructed. Therefore, we propose an
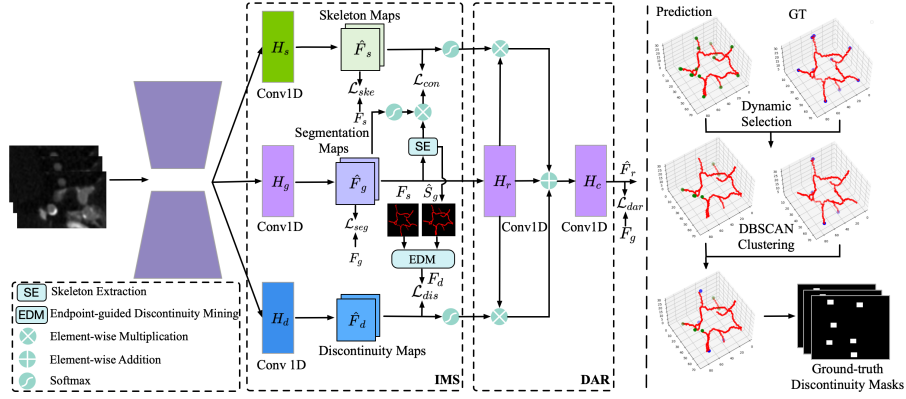
Fig. 2. An overview of the proposed GLCP. **Left)** The input patch is first fed into an encoder-decoder backbone to extract image representations. Next, an IMS module uses the encoded representations as input to produce segmentation, discontinuity and skeleton maps, based on which we design a DAR module to integrate global skeleton attention with local discontinuity attention for refining segmentation results. **Right)** In EDM, endpoints indicating potential discontinuities are identified and used to form the ground-truth discontinuity masks. Note that SE and EDM are only used for training.

endpoint-guided discontinuity mining strategy to generate ground-truth discontinuity masks. In more detail, we first extract skeletons (F_s and $\hat{S}_g$) from the corresponding ground truth F_g and predicted segmentation masks $\hat{F}_g$, respectively by skeleton extraction algorithms [20,22]. Given $\hat{S}_g$ and F_s, we aim to identify potential structural fragmentation by detecting and analyzing their endpoints. In this way, we first detect all endpoints in $\hat{S}_g$ and F_s through a convolution operation with a fixed kernel, resulting in the sets $\hat{P}_g = \{\hat{p}_1, \hat{p}_2, ..., \hat{p}_M\}$ and $P_g = \{p_1, p_2, ..., p_N\}$, respectively, where M and N are the number of endpoints. To identify potential discontinuity regions, we then propose a dynamic endpoint selection strategy, which calculates distances between the endpoints in $\hat{P}_g$ and P_g, and adaptively selects discontinuity points based on statistical thresholds. Formally, for each $\hat{p}_i \in \hat{P}_g$, we calculate the shortest Euclidean distance to all the points in P_g:

$$\hat{d}_i = \min_{p_j \in P_g} \|\hat{p}_i - p_j\|_2, \tag{1}$$

where $\hat{d}_i$ represents the shortest distance from $\hat{p}_i$ to P_g. Then the distances for all endpoints in $\hat{P}_g$ form the set $\hat{D}_g = \{\hat{d}_1, \hat{d}_2, ..., \hat{d}_M\}$.

To adaptively determine discontinuity points, we define a dynamic threshold $\hat{\tau}$ based on the mean and standard deviation of $\hat{D}_g$. Any point $\hat{p}_i \in \hat{P}_g$ with $\hat{d}_i > \hat{\tau}$ is then identified as a potential discontinuity point, forming a subset:

$$\widetilde{P}_g = \{\hat{p}_i \in \hat{P}_g \mid \hat{d}_i > \hat{\tau}\}, \hat{\tau} = \text{mean}(\hat{D}_g) + \text{std}(\hat{D}_g) \tag{2}$$

where mean($\cdot$) and std($\cdot$) represent the mean and standard deviation, respectively. Typically, the selected $\widetilde{P}g$ includes endpoints that do not exist in the GT but are introduced in the prediction due to discontinuities. However, endpoints present in the GT may also be lost in the prediction. Considering these regions, we also select critical positions from the GT to supplement $\widetilde{P}g$. Similarly, we can obtain the set $\overline{P}_g$ of discontinuity points from P_g using Eqs. (1) and (2). The two sets of discontinuity points are then merged into a unified set $P' = \widetilde{P}_g \cup \overline{P}_g$. To further refine the selection of discontinuity points, we apply DBSCAN clustering [10] to P' to obtain a new set P, which groups nearby points within a distance, effectively reducing redundancy caused by closely spaced discontinuity points. We randomly select one point from these neighboring points during the training to balance the training data and prevent over-representation of densely clustered discontinuity regions.

Directly predicting sparse and isolated discontinuity points is challenging due to their irregular distribution. Instead, we construct ground-truth discontinuity masks by expanding the identified discontinuity points to their local neighbourhoods. For each discontinuity point $p_i \in P$, we define a cube window R_i of size $w_z \times w_y \times w_x$, centered on p_i, which is expressed as:

$$R_i = \{(x, y, z) \mid |x - p_i^x| \leq w_x/2, |y - p_i^y| \leq w_y/2, |y - p_i^z| \leq w_z/2\} \quad (3)$$

where p_i^x, p_i^y and p_i^z are the coordinates of p_i. Each pixel within this region is assigned with a value of 1, indicating the discontinuity area. The final ground-truth discontinuity mask F_f is obtained by taking the union of all such regions, denoted as $F_d = \bigcup_{p_i \in P} R_i$.

Inspired by [7], we also incorporate a skeleton prediction task to complement the discontinuity prediction task. Unlike [7] using an additional decoder, we design a skeleton head H_s to predict skeleton maps $\hat{F}_s$, with the corresponding GT denoted as F_s. Additionally, we impose self-supervised consistency constraints between global skeletons from this task and the main objective, respectively, to promote their interaction, thus allowing for more accurate and consistent predictions. The consistency loss $\mathcal{L}_{con}$ can be defined as:

$$\mathcal{L}_{con} = \text{KL}\left(\sigma(\hat{F}_g) \otimes \hat{S}_g, \psi(\hat{F}_s)\right) + \text{KL}\left(\hat{F}_s, \psi(\sigma(\hat{F}_g) \otimes \hat{S}_g)\right) \quad (4)$$

where $\sigma(\cdot)$ represents softmax activation function, $\psi(\cdot)$ denotes a function that truncates the gradient of input, and $KL(\cdot)$ is the Kullback-Leibler divergence loss. The proposed consistency loss encourages the two tasks to focus on the same global topological structure by aligning the probability distributions of the skeletons, which promotes stronger inter-task interaction. The use of the gradient truncation function $\psi(\cdot)$ stabilizes the training process by preventing direct interference between the two tasks while still maintaining consistency.

2.2 Dual-Attention-based Refinement (DAR)

The proposed IMS encourage the model to better capture global and local structural properties by jointly learning segmentation (e.g., $\hat{F}_g$), skeleton (i.e., $\hat{F}_s$) and

discontinuity maps (e.g., $\hat{F}_d$). To further improve the segmentation quality, a DAR module is proposed to refine the initial prediction results, e.g., $\hat{F}_g$. Considering the global connectivity and local fragmentation indicated by the skeleton and discontinuity maps, we explicitly construct global and local attention maps to guide the refinement process. Hence, the refined segmentation maps $\hat{F}_r$ are formulated as:

$$\hat{F}_r = H_c\big(\sigma(\hat{F}_s) \otimes H_r(\hat{F}_g) + \sigma(\hat{F}_d) \otimes H_r(\hat{F}_g) + H_r(\hat{F}_g)\big) \tag{5}$$

where $\sigma(\cdot)$ is used to yield skeleton and discontinuity attention maps. $H_r(\cdot)$ and $H_c(\cdot)$ denote convolution operations with the kernel size of 1, where $H_r(\cdot)$ maps $\hat{F}_g$ to a higher-dimensional space, and $H_c(\cdot)$ generates the final prediction results.

By integrating the DAR with IMS, the overall loss function is calculated as:

$$\mathcal{L}_{total} = \mathcal{L}_{ims} + \alpha\mathcal{L}_{con} + \beta\mathcal{L}_{dar} \tag{6}$$

where α and β are weight factors. $\mathcal{L}_{ims}$ represents the sum of the losses of the segmentation ($\mathcal{L}_{seg}$), discontinuity($\mathcal{L}_{dis}$) and skeleton ($\mathcal{L}_{ske}$) predictions. $\mathcal{L}_{dar}$ denotes the refinement loss. We use cross-entropy (CE) loss for $\mathcal{L}_{dar}$, while $\mathcal{L}_{seg}$, $\mathcal{L}_{dis}$ and $\mathcal{L}_{ske}$ follow the default loss functions in nnUNet [9].

3   Experiments

3.1   Datasets and Evaluation Metrics

We evaluate our proposed method on three benchmark datasets. The STARE dataset [5] is used for 2D retinal vessel segmentation, containing 10 images for training and 10 for testing. The CCA dataset [25] provides 20 CTA images depicting coronary artery disease, with 16 for training and 4 for testing. The MICCAI 2023 TopCoW Challenge dataset [24] consists of 90 brain CTA cases, with 72 for training and 18 for testing. Both binary and multi-class settings on TopCoW are employed in this work. We report the volumetric scores (Dice and clDice [20]), topology errors (Betti Error [6] β for the sum of Betti Numbers β_0 and β_1) and distance errors (Hausdorff Distance (HD) [21]).

3.2   Implementation Details

Similar to [19], nnUNet V2 is selected as our baseline. The proposed GLCP can be seamlessly integrated with the baseline by adding only four lightweight convolution layers (H_d, H_s, H_r and H_c) at the end, without introducing significant modifications. R_i is set to one-eighth of the input patch size, and both α and β are set to 0.5. For other training and optimization settings, we follow the default configurations of nnUNet for each dataset. In the skeleton prediction task of the multi-class setting, all classes except the background are treated as the foreground object to extract the overall skeleton.

Table 1. Comparison with the state-of-the-art methods on the STARE, ToPCoW-binary and CCA datasets. * means training with default CE and Dice loss. MD means multi-decoder configuration [7] consisting of skeleton and edge prediction tasks.

Method	STARE				ToPCoW-binary				CCA			
	Dice↑	clDice↑	β↓	HD↓	Dice↑	clDice↑	β↓	HD↓	Dice↑	clDice↑	β↓	HD↓
nnUNet* [9]	82.92	86.25	4.60	6.94	90.43	95.41	1.94	1.68	86.73	87.53	35.75	21.30
+clDice [20]	83.30	86.99	3.90	5.30	90.63	95.57	1.72	1.63	87.57	88.19	19.50	19.27
+cbDice [19]	83.05	86.34	4.50	5.17	90.41	95.44	1.67	1.65	86.74	87.62	17.25	26.53
+ ske-recall [11]	83.39	87.11	3.70	5.06	91.02	95.44	1.56	**1.58**	87.84	89.50	15.00	16.42
+ Ours (GLCP)	**83.67**	**87.44**	**3.00**	**4.64**	**91.34**	**95.58**	**1.17**	1.60	**87.94**	**90.55**	**10.50**	**14.03**
+ MD [7]	83.26	87.02	4.10	5.28	90.84	95.43	1.61	1.61	87.32	88.01	25.75	16.69
+ Ours (IMS)	**83.58**	**87.42**	**3.20**	**4.74**	**91.24**	**95.52**	**1.39**	**1.53**	**88.09**	**89.95**	**12.25**	**14.26**
SwinUNETR* [4]	82.16	86.12	5.10	7.21	90.15	95.42	2.00	1.84	85.44	86.45	44.25	26.34
+ Ours (GLCP)	**83.25**	**86.96**	**3.60**	**4.96**	**91.27**	**95.52**	**1.24**	**1.64**	**87.12**	**89.63**	**17.50**	**20.25**

Table 2. Comparison with the state-of-the-art methods on ToPCoW-multi.

Method	Dice↑	clDice↑	β↓	HD↓
nnUNet* [9]	71.15	86.84	0.57	3.21
+clDice [20]	71.41	88.62	0.50	3.02
+cbDice [19]	71.55	88.49	0.48	3.14
+ ske-recall [11]	72.26	88.63	0.49	2.70
+ Ours (GLCP)	**74.22**	**89.15**	**0.39**	**2.64**
+ MD [7]	72.12	88.83	0.51	2.82
+ Ours (IMS)	**73.33**	**89.03**	**0.41**	**2.79**
SwinUNETR* [4]	70.12	86.13	0.63	3.30
+ Ours (GLCP)	**72.45**	**88.54**	**0.44**	**2.84**

Table 3. Ablation study of components.

Dataset	Ske	Dis	Self	DAR	Dice↑	clDice↑	β↓	HD↓
STARE					82.92	86.25	4.60	6.94
	✓				83.16	86.31	4.40	6.22
		✓			83.54	86.72	4.10	4.89
	✓	✓			83.45	87.08	3.60	5.62
	✓	✓	✓		83.58	87.42	3.20	4.74
	✓	✓	✓	✓	**83.67**	**87.44**	**3.00**	**4.64**
ToPCoW -multi					71.15	86.84	0.57	3.21
	✓				71.59	88.50	0.54	3.52
		✓			72.38	88.78	0.49	2.82
	✓	✓			73.26	**89.57**	0.43	2.85
	✓	✓	✓		73.33	89.03	0.41	2.79
	✓	✓	✓	✓	**74.22**	89.15	**0.39**	**2.64**

3.3 Comparison Results and Ablation Study

Table 1 presents the quantitative results of our method compared to five other SOTA loss functions. We use the default loss weights provided in their publicly available codes to implement these methods. For binary segmentation tasks on STARE, ToP-CoW and CCA, our proposed GLCP (IMS+DAR) achieves improvement in terms of Dice and clDice scores, compared to the closest competitors. Meanwhile, from a topological perspective, our GLCP demonstrates superior topological continuity, achieving the lowest β errors of 3.00, 1.17, and 10.50, respectively. These results highlight the ability of our method to effectively capture both global and local structural characteristics of 2D and 3D tubular networks, delivering more accurate segmentation performance and ensuring a more continuous topology. In addition, our method achieves competitive results in the quality of boundary extraction, attaining the lowest HD on both the STARE and CCA datasets. We also compare our IMS with a similar multi-task learning framework proposed in [7]. For a fair comparison, we re-implement the multi-decoder (MD) prediction network from [7] based on the nnUNet framework. The results demonstrate that our IMS significantly outperforms MD across all metrics. In addition to binary segmentation tasks, we also evaluated our method on
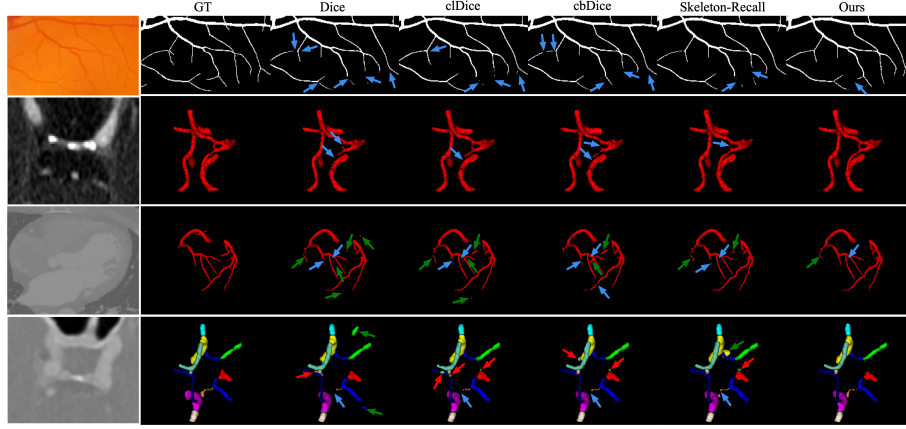
Fig. 3. Qualitative results on the STARE, ToPCoW-binary, CCA and ToPCoW-multi datasets (listed from top to bottom). Blue, green, and red arrows indicate regions of false negatives, false positives, and misclassifications, respectively, highlighting the challenges in segmentation and the improvements achieved by our method.

multi-class segmentation tasks, as shown in Table 2. Our GLCP achieves the best Dice of 74.22% and clDice of 89.15% among all the methods in comparison, representing improvements of 3.07% and 2.31% over the baseline, respectively. In terms of topology and distance errors, our method also outperforms these approaches, demonstrating superior topological continuity and boundary accuracy. Finally, to validate the generalizability of our method, we apply GLCP to the transformer-based SwinUNETR [4], as shown in the last two rows of Tables 1 and 2, demonstrating its effectiveness across different architectures.

To better illustrate the effectiveness of our method, we showcase several qualitative results on both 2D and 3D examples in Fig. 3. The results demonstrate that our method is more effective than other approaches in eliminating topology errors and misclassifications when compared to the ground truth.

Table 3 summarizes the ablation study results on the STARE and ToPCoW-multi datasets, highlighting the effectiveness of each proposed component. Specifically, we can observe that the discontinuity prediction task (Dis) plays a more important role in correcting topology errors than the skeleton prediction task (Ske) and combining them can gain further improvement. Besides, incorporating the self-supervised (self) consistency constraints loss achieves slight improvements across nearly all the metrics. This can be attributed to the ability of this loss to enhance inter-task interactions by boosting the alignment between the skeletons from the main task and the skeleton prediction task. Finally, the DAR module contributes to further improvements by refining the segmentation maps based on the proposed dual-attention mechanism, achieving the best performance across Dice, β error and HD.

4 Conclusions

In this paper, we proposed a novel GLCP framework to address the challenge of structural fragmentation in tubular structure segmentation, particularly for vascular networks. We introduced IMS to jointly perceive global topological structures and local discontinuity regions, ensuring improved segmentation accuracy and continuity. Additionally, DAR is proposed to refine segmentation quality through the dual-attention mechanism. Extensive experiments on 2D and 3D datasets demonstrate that our method achieves state-of-the-art performance. Future research will explore the generalizability of our approach across additional segmentation networks and validate its effectiveness on larger and more complex vascular datasets.

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

References

1. Alirr, O.I., Abd Rahni, A.A.: An automated liver vasculature segmentation from ct scans for hepatic surgical planning. International Journal of Integrated Engineering **13**(1), 188–200 (2021)
2. Clough, J.R., Byrne, N., Oksuz, I., Zimmer, V.A., Schnabel, J.A., King, A.P.: A topological loss function for deep-learning based image segmentation using persistent homology. IEEE transactions on pattern analysis and machine intelligence **44**(12), 8766–8778 (2020)
3. Dong, S., Pan, Z., Fu, Y., Yang, Q., Gao, Y., Yu, T., Shi, Y., Zhuo, C.: Deu-net 2.0: Enhanced deformable u-net for 3d cardiac cine mri segmentation. Medical Image Analysis **78**, 102389 (2022)
4. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In: International MICCAI brainlesion workshop. pp. 272–284. Springer (2021)
5. Hoover, A., Goldbaum, M.: Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. IEEE transactions on medical imaging **22**(8), 951–958 (2003)
6. Hu, X., Li, F., Samaras, D., Chen, C.: Topology-preserving deep image segmentation. Advances in neural information processing systems **32** (2019)
7. Huang, J., Zhou, Y., Luo, Y., Liu, G., Guo, H., Yang, G.: Representing topological self-similarity using fractal feature maps for accurate segmentation of tubular structures. In: European Conference on Computer Vision. pp. 143–160. Springer (2025)
8. Ii, S., Kitade, H., Ishida, S., Imai, Y., Watanabe, Y., Wada, S.: Multiscale modeling of human cerebrovasculature: A hybrid approach using image-based geometry and a mathematical algorithm. PLoS computational biology **16**(6), e1007943 (2020)

9. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)

10. Khan, K., Rehman, S.U., Aziz, K., Fong, S., Sarasvady, S.: Dbscan: Past, present and future. In: The fifth international conference on the applications of digital information and web technologies (ICADIWT 2014). pp. 232–238. IEEE (2014)

11. Kirchhoff, Y., Rokuss, M.R., Roy, S., Kovacs, B., Ulrich, C., Wald, T., Zenk, M., Vollmuth, P., Kleesiek, J., Isensee, F., Maier-Hein, K.: Skeleton recall loss for connectivity conserving and resource efficient segmentation of thin tubular structures. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T., Varol, G. (eds.) Computer Vision – ECCV 2024. pp. 218–234. Springer Nature Switzerland, Cham (2024)

12. Li, L., Ma, Q., Ouyang, C., Li, Z., Meng, Q., Zhang, W., Qiao, M., Kyriakopoulou, V., Hajnal, J.V., Rueckert, D., et al.: Robust segmentation via topology violation detection and feature synthesis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 67–77. Springer (2023)

13. Li, Y., Zhang, Y., Liu, J.Y., Wang, K., Zhang, K., Zhang, G.S., Liao, X.F., Yang, G.: Global transformer and dual local attention network via deep-shallow hierarchical feature fusion for retinal vessel segmentation. IEEE Transactions on Cybernetics **53**(9), 5826–5839 (2022)

14. Liu, C., Ma, B., Ban, X., Xie, Y., Wang, H., Xue, W., Ma, J., Xu, K.: Enhancing boundary segmentation for topological accuracy with skeleton-based methods. arXiv preprint arXiv:2404.18539 (2024)

15. Ma, Y., Hao, H., Xie, J., Fu, H., Zhang, J., Yang, J., Wang, Z., Liu, J., Zheng, Y., Zhao, Y.: Rose: a retinal oct-angiography vessel segmentation dataset and new model. IEEE transactions on medical imaging **40**(3), 928–939 (2020)

16. Mou, L., Zhao, Y., Chen, L., Cheng, J., Gu, Z., Hao, H., Qi, H., Zheng, Y., Frangi, A., Liu, J.: Cs-net: Channel and spatial attention network for curvilinear structure segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22. pp. 721–730. Springer (2019)

17. Qi, X., Yang, G., He, Y., Liu, W., Islam, A., Li, S.: Contrastive re-localization and history distillation in federated cmr segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 256–265. Springer (2022)

18. Qi, Y., He, Y., Qi, X., Zhang, Y., Yang, G.: Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 6070–6079 (2023)

19. Shi, P., Hu, J., Yang, Y., Gao, Z., Liu, W., Ma, T.: Centerline boundary dice loss for vascular segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 46–56. Springer (2024)

20. 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)

21. Taha, A.A., Hanbury, A.: Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. BMC medical imaging **15**, 1–28 (2015)

22. Van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T.: scikit-image: image processing in python. PeerJ **2**, e453 (2014)

23. Wang, Y., Wei, X., Liu, F., Chen, J., Zhou, Y., Shen, W., Fishman, E.K., Yuille, A.L.: Deep distance transform for tubular structure segmentation in ct scans. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3833–3842 (2020)
24. Yang, K., Musio, F., Ma, Y., Juchler, N., Paetzold, J.C., Al-Maskari, R., Höher, L., Li, H.B., Hamamci, I.E., Sekuboyina, A., et al.: Benchmarking the cow with the topcow challenge: Topology-aware anatomical segmentation of the circle of willis for cta and mra. ArXiv pp. arXiv–2312 (2024)
25. Yang, X., Xu, L., Yu, S., Xia, Q., Li, H., Zhang, S.: Segmentation and vascular vectorization for coronary artery by geometry-based cascaded neural network. IEEE Transactions on Medical Imaging (2024)
26. Yang, X., Li, Z., Guo, Y., Zhou, D.: Dcu-net: A deformable convolutional neural network based on cascade u-net for retinal vessel segmentation. Multimedia Tools and Applications **81**(11), 15593–15607 (2022)
27. Zhang, X., Zhang, J., Ma, L., Xue, P., Hu, Y., Wu, D., Zhan, Y., Feng, J., Shen, D.: Progressive deep segmentation of coronary artery via hierarchical topology learning. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 391–400. Springer (2022)
28. Zhao, H., Li, H., Cheng, L.: Improving retinal vessel segmentation with joint local loss by matting. Pattern Recognition **98**, 107068 (2020)