Supplementary Material for the Paper "DeformCL: Learning Deformable Centerline Representation for Vessel Extraction in 3D Medical Image"

1. Dataset Details

We use four datasets to evaluate our proposed framework. For the **HaN-Seg** [7] dataset, we randomly split the publicly available 42 CT scans into 31 and 11 images as train and test data, and the MR images were not used. Left and right carotid arteries are utilized as segmentation targets. In the case of the **HNCTA** dataset containing 358 head and neck CT scans, clinical experts provide accurate segmentation labels for nine categories of vessels, including the left and right vertebral arteries (VA), internal carotid arteries (ICA), common carotid arteries (CCA), posterior communicating arteries (PoCA), and basilar artery (BA). The dataset is randomly split into 242, 58, and 58 images as train, validation, and test data. For the **ASOCA** [1, 2] dataset, we split the dataset randomly into 33 training and 7 testing images. For the **ImageCAS** [13] dataset, we fulfill the vessel category annotations for 350 cases with experienced radiologists and split it into 235, 50, and 65 for train, validation, and test data. We evaluate the proposed method on the three main coronary arteries (LAD, LCX, and RCA) for all CCTA datasets including ASOCA and ImageCAS.

2. Additional Implementation Details

During the training process, we employ RandomCrop and RandomFlip as data augmentations to mitigate overfitting, with consistent settings applied across all methods. For larger datasets such as ImageCAS [13] and HNCTA, we train our model and all baselines for 12,000 iterations on 8 NVIDIA GeForce RTX 3090 GPUs. For smaller datasets like ASOCA [1, 2] and HaN-Seg [7], we train our model and all baselines for 4,000 iterations. The batch size is set to 8 by default for all experiments. We utilize the AdamW Optimizer [5] with a base learning rate of 1×10^{-3}. Additionally, we employ a cosine learning rate schedule with warmup. All models are implemented based on PyTorch [6].

We provide the required resources and runtime analusis on 3D UNet in Table 1, and we can conclude that the increase of DeformCL in resource usage is minimal.

Table 1. Computational costs and runtime analysis.

Method	Params (M)	FLOPs (G)	GPU Memory (MB)	Inference Time (s/img)
SoftDice	2.08	3.11	3046	0.12
Ours(SoftDice)	2.20	4.03	3228	0.18

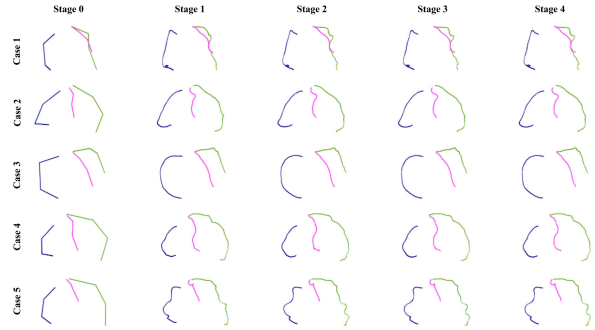

Figure 1. Visualizations of the centerline deformation process on the ImageCAS dataset. Each row corresponds to a different case, and each column represents a different deformation stage. The first subfigure in each row (Stage 0) displays the centerline template generated using linear interpolation, while subsequent subfigures show the centerline predictions after deformation stages 1, 2, 3, and 4, respectively.

3. Analysis of Centerline Deformation Process

We present visualizations of the centerline deformation process on the ImageCAS dataset [13], aimed at enhancing understanding of DeformCL, as depicted in Figure 1. Stage 0 represents the centerline templates generated by linear interpolation on control points. It is observed that the core structure of the centerlines typically stabilizes within the initial one or two stages, with subsequent stages primarily focused on refining finer details. Furthermore, we note that the predicted centerlines naturally possess sufficient smoothness, which proves beneficial for the subsequent clinical tasks such as CPR [3] image generation.

4. Additional Experimental Results

For broader comparisons, we include Transformer backbones Swin UNETR in Table 2. We also provide more ablation studies as follows.

Table 2. Results of segmentation and centerline on HaN-Seg dataset with Swin UNETR as backbone.

Backbone	Method	Centerline Scores		Volumetric Scores (%) ↑		Topological Error ↓			Distance ↓
		F1 (%) ↑	Chamfer ↓	Dice	clDice	β_0	β_1	Euler	HD95
Swin UNETR	softDice	91.72	11.37	77.05	89.06	1.815	0.206	1.738	4.624
Swin UNETR	**Ours** (softDice)	94.54	7.53	79.45	89.46	1.410	0.051	1.388	3.701
Swin UNETR	clDice	92.79	15.67	78.41	89.50	2.116	0.167	1.977	4.513
Swin UNETR	**Ours** (clDice)	94.11	7.10	79.29	89.62	1.982	0.096	1.905	3.718

Table 3. Ablation on curvilinear interpolation method on HaN-Seg dataset.

Interpolation Methods	Volumetric Scores (%) ↑	
	Dice	clDice
B-Spline, 2-order	**79.95**	88.20
B-Spline, 3-order	79.53	88.10
Linear (Default)	79.58	**88.86**

Curvilinear Interpolation Method. During adaptive template generation, curvilinear interpolation is essential for generating the initial centerline template based on control points. In this subsection, we investigate the efficacy of different interpolation methods, encompassing second- and third-order B-Spline Interpolation, as well as Linear Interpolation, which serves as our default setting owing to its simplicity. Figure 2 provides an illustration of the aforementioned interpolation methods. Experimental results in Table 3 demonstrate that different interpolation methods yield similar performance in vessel segmentation. This suggests that the overall framework is robust to initial templates, and several deformation stages are adequate to deform the centerline templates towards the ground truth centerlines effectively. This observation aligns with the analysis in Section 3 – after the initial one or two stages, the core structure of centerlines will be stabilized.

Interaction Approach. As discussed in the main paper, previous methods [8, 9, 11, 12] lack an effective approach for curvilinear feature aggregation due to their discrete representations. In contrast, DeformCL inherently exhibits a graph structure, enabling point feature interaction along the tubular curve efficiently. Table 4 compares different interaction approaches, including Graph Convolutional Network (GCN) [4] and Transformer [10]. The results indicate that the improvement with Transformer is more significant than with GCN, possibly due to its powerful long-distance modeling capability.

Figure 2. Visualization of centerline templates obtained through different interpolation methods: linear interpolation, second-order B-Spline, third-order B-Spline.

Table 4. Ablation on interaction approach on HaN-Seg dataset.

Interaction Approaches	Volumetric Scores (%) ↑	
	Dice	clDice
∅	78.75	88.40
GCN	78.77	88.61
Transformer (Default)	**79.58**	**88.86**

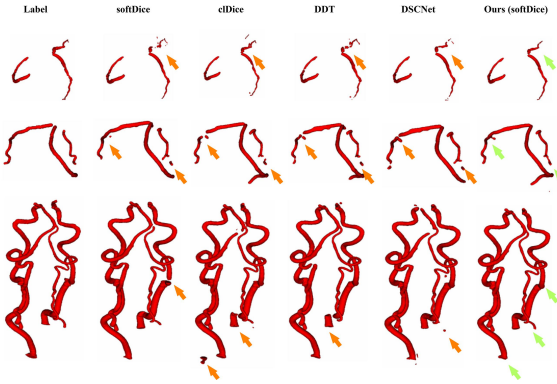

Figure 3. Additional visualizations of the predictions from different methods on 3D medical image datasets. The orange arrows point out the areas where previous methods exhibit errors, while the green arrows point out the improved results from our approach.

5. Additional Qualitative Results

In this section, we provide additional qualitative results including the visual comparison of segmentation results and SCPR images in Figure 3 and Figure 4.

Moreover, we present a failure case involving almost completely occluded vessels, which are highly challenging for deep learning models due to their near invisibility. As shown in Figure 5, our method produces fragmented predictions. However, it outperforms the previous mask-based representation, thanks to its continuous property.

References

[1] Ramtin Gharleghi, Dona Adikari, Katy Ellenberger, Sze-Yuan Ooi, Chris Ellis, Chung-Ming Chen, Ruochen Gao, Yuting He, Raabid Hussain, Chia-Yen Lee, Jun Li, Jun Ma, Ziwei Nie, Bruno Oliveira, Yaolei Qi, Youssef Skandarani, João L. Vilaça, Xiyue Wang, Sen Yang, Arcot Sowmya, and Susann Beier. Automated segmentation of normal and diseased coronary arteries – the asoca challenge. *Computerized Medical Imaging and Graphics*, 97:102049, 2022. 1

[2] R Gharleghi, D Adikari, K Ellenberger, M Webster, C Ellis, A Sowmya, S Ooi, and S Beier. Annotated computed tomography coronary angiogram images and associated data of normal and diseased arteries. *Scientific Data*, 10(1):128, 2023. 1

[3] Armin Kanitsar, Dominik Fleischmann, Rainer Wegenkittl, Petr Felkel, and Eduard Groller. *CPR-curved planar reformation*. IEEE, 2002. 1

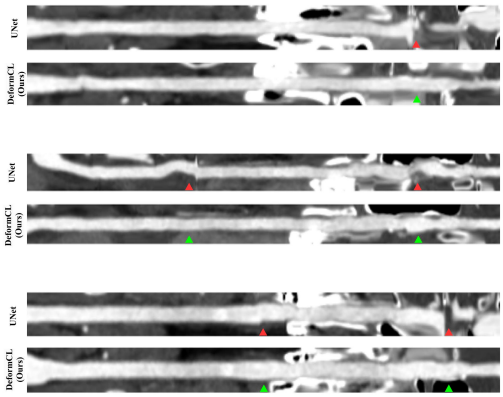

Figure 4. Additional visual comparison of SCPR images reconstructed from the predictions of UNet and DeformCL. Red triangles indicate the regions where the vessel predictions by UNet are not correct, resulting in severe errors on SCPR images, while green ones indicate the improved results from our DeformCL.

Figure 5. Performance of centerlines on occluded vessels.

[4] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016. 2

[5] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017. 1

[6] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, and et al. Zeming Lin. Pytorch: An imperative style, high-performance deep learning library. In *NeurIPS*, 2019. 1

[7] Gašper Podobnik, Primož Strojan, Primož Peterlin, Bulat Ibragimov, and Tomaž Vrtovec. Han-seg: The head and neck organ-at-risk ct and mr segmentation dataset. *Medical physics*, 50(3):1917–1927, 2023. 1

[8] Yaolei Qi, Yuting He, Xiaoming Qi, Yuan Zhang, and Guanyu Yang. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6070–6079, 2023. 2

[9] Suprosanna Shit, Johannes C Paetzold, Anjany Sekuboyina, Ivan Ezhov, Alexander Unger, Andrey Zhylka, Josien PW Pluim, Ulrich Bauer, and Bjoern H Menze. cldice-a novel topology-preserving loss function for tubular structure segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16560–16569, 2021. 2

[10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. 2

[11] Dong Wang, Zhao Zhang, Ziwei Zhao, Yuhang Liu, Yihong Chen, and Liwei Wang. Pointscatter: Point set representation for tubular structure extraction. In *European Conference on Computer Vision*, pages 366–383. Springer, 2022. 2

[12] Yan Wang, Xu Wei, Fengze Liu, Jieneng Chen, Yuyin Zhou, Wei Shen, Elliot K Fishman, and Alan L Yuille. Deep distance transform for tubular structure segmentation in ct scans. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3833–3842, 2020. 2

[13] An Zeng, Chunbiao Wu, Guisen Lin, Wen Xie, Jin Hong, Meiping Huang, Jian Zhuang, Shanshan Bi, Dan Pan, Najeeb Ullah, et al. Imagecas: A large-scale dataset and benchmark for coronary artery segmentation based on computed tomography angiography images. *Computerized Medical Imaging and Graphics*, 109:102287, 2023. 1