Topological-Aware Regularization for Semi-Supervised Intracranial Aneurysm Vessel Segmentation

Feiyang Xiao[1,2,*] FYXIAO24@M.FUDAN.EDU.CN
Yichi Zhang[1,2,*] ZHANGYICHI23@M.FUDAN.EDU.CN
Xigui Li[1,2] LIXIGUI@FUDAN.EDU.CN
Yuanye Zhou[2,3] ZHYY2009@163.COM
Chen Jiang[1,2] JIANGCHEN@SAIS.COM.CN
Xin Guo[1,2] GUOXIN@SAIS.COM.CN
Limei Han[1,2] HANLIMEI@FUDAN.EDU.CN
Yuxin Li[4,†] LIYUXIN@FUDAN.EDU.CN
Fengping Zhu[4,†] ZHUFENGPING@FUDAN.EDU.CN
Yuan Cheng[1,2,†] CHENG_YUAN@FUDAN.EDU.CN
[1] *Artificial Intelligence Innovation and Incubation Institute, Fudan University, Shanghai, China*
[2] *Shanghai Academy of Artificial Intelligence for Science, Shanghai, China*
[3] *Hong Kong Polytechnic University, Hong Kong*
[4] *Huashan Hospital, Fudan University, Shanghai, China*

Editors: Under Review for MIDL 2026

Abstract

Accurate segmentation of intracranial aneurysm and their parent vessels (IA-Vessel) from magnetic resonance angiography is a critical prerequisite for computational fluid dynamics-based rupture risk assessment. While deep learning methods can automate this laborious task, they are hindered by the high cost and scarcity of expert annotations. Most existing semi-supervised methods focus on enforcing regional constraints while largely ignoring topological constraints, which is insensitive to subtle but critical errors like vessel adhesion or surface irregularities, which are often unsuitable for downstream applications. To address this gap, we introduce topological-aware regularization (TAR) for by incorporating the learning of local vascular topology to ensure the precise and geometrically correct segmentation of the IA-Vessel complex using only a small amount of labeled data. Experimental results on a multi-center MRA dataset show that our framework efficiently utilizes unlabeled data and outperforms state-of-the-art semi-supervised segmentation methods. Instead of being restricted to a fixed framework, TAR is a plug-and-play strategy that can be seamlessly integrated into various semi-supervised frameworks to further boost their performance. The code and model weights will be made publicly available after the paper is accepted.

Keywords: Intracranial Aneurysm Segmentation, Semi-Supervised Learning, Topological-Aware Regularization

[*] Contributed equally
[†] Corresponding authors.

1. Introduction

Intracranial aneurysm (IA) is a pathological dilation of blood vessels, primarily occurring at arterial bifurcations (Schievink, 1997). Although often initially asymptomatic, IAs can enlarge and rupture, leading to severe morbidity and mortality (Cebral et al., 2005). The accurate assessment of rupture risk is therefore essential for clinical intervention (Etminan and Rinkel, 2016). Computational Fluid Dynamics (CFD) offers key biomechanical insights for this assessment by quantifying hemodynamic parameters and has been widely applied in biomedical research (Li et al., 2025; Morris et al., 2016; Wang et al., 2025).

Magnetic resonance angiography (MRA) serves as a high-resolution, non-invasive imaging modality for visualizing the detailed anatomical features of aneurysms (Pierot et al., 2013). An accurate segmentation of the intracranial aneurysm and parent vessels (IA-Vessel) from these images is a critical step for performing CFD analysis (Patel et al., 2023). As manual segmentation remains a labor-intensive and time-consuming procedure for radiologists, developing automated methods is highly desirable. Deep learning has emerged as a state-of-the-art approach, achieving results, though its development depends on the availability of large amount of annotated data (Antonelli et al., 2022; Ma et al., 2022; Qu et al., 2023). However, acquiring extensively annotated medical datasets is challenging because the process requires domain expertise and is consequently expensive and resource-intensive (Tajbakhsh et al., 2020; Shi et al., 2024). Given the abundance of unlabeled data, semi-supervised learning presents an attractive solution by leveraging a limited set of labeled data (Jiao et al., 2023).

Existing semi-supervised medical image segmentation approaches can be broadly categorized into pseudo labeling and unsupervised regularization. Pseudo labeling methods focus on generating pseudo labels for unlabeled data using either pre-trained models or dynamically updated networks, and then treating these predictions as weak annotations for further training (Thompson et al., 2022; Yao et al., 2022). To mitigate the noise introduced by incorrect pseudo labels, various quality-control mechanisms have been developed (Li et al., 2020; Lu et al., 2023). Unsupervised regularization avoids explicit pseudo-label generation and instead leverages unlabeled data by enforcing invariance constraints to learn robust representations (Zhang et al., 2023; Luo et al., 2022). A representative technique within this category is consistency learning, which encourages stable predictions under different perturbations of the same input, as exemplified by the Mean Teacher framework (Tarvainen and Valpola, 2017) and its variants (Yu et al., 2019; Shi et al., 2023; Assefa et al., 2025). Several recent methods also explore leveraging ptr-trained foundation models like SAM (Zhang et al., 2024a) to enhance the stability of semi-supervised learning with limited annotation (Zhang et al., 2024b, 2025). Nevertheless, a crucial limitation persists as most existing models are evaluated on region overlap-based metrics, which is insensitive to geometric topological abnormalities such as vessel adhesion and surface irregularities. This often results in segmentation outcomes that, despite high Dice scores, are unsuitable for downstream CFD application due to mesh generation failure or flow field distortion (Xiao et al., 2025).

To overcome these challenges, we introduce topological-aware regularization (TAR) for semi-supervised intracranial aneurysm vessel segmentation. By incorporating the learning of local vascular topology, our method ensures the precise segmentation of the IA-Vessel complex using a small amount of labeled data. Experimental results on a large-scale multi-center MRA dataset show that our framework efficiently utilize unlabeled data and outperform state-of-the-art semi-supervised segmentation methods with better performance. Notably, instead of being restricted to a fixed framework, TAR is a plug-and-play strategy that can be seamlessly integrated into various semi-supervised frameworks to further boost their performance by moving beyond improving segmentation accuracy alone to guarantee the topological integrity of the output.

2. Method

2.1. Task Definition

To ease the description of methodology, we can define the task of semi-supervised medical image segmentation as follows. Given a training dataset D, it is split into a labeled set with M cases, denoted as $D_L = \{x_i^l, y_i\}_{i=1}^{M}$, and an unlabeled set with N cases, denoted as $D_U = \{x_i^u\}_{i=1}^{N}$. Here, x^l and x^u represent the input images, and y_i is the corresponding ground-truth segmentation for the labeled data. The model is required to utilize both D_L and D_U during the training phase, enabling the network to produce segmentation results for new images during inference that are comparable to those of an optimal model trained on a fully labeled dataset.

To accomplish this, semi-supervised learning is typically designed as a two-fold task. First, a supervised loss is applied to the labeled set D_L similar to fully-supervised methods to ensure the network effectively learns features from the available labels. Second, an unsupervised regularization term is introduced for the unlabeled set D_U. For example, consistency regularization aims to penalize differences in predictions for the same input under various perturbations. By doing so, it forces the network to maintain stable predictions against disturbances in the input space, which in turn smoothly propagates label information from labeled to unlabeled regions.

2.2. Semi-Supervised Backbone

The Mean Teacher framework (Tarvainen and Valpola, 2017) is widely used in semi-supervised image segmentation. It consists of a student model and a teacher model, which share an identical structure but employ different parameter update strategies. During training, labeled data is fed into the student model, and a supervised loss is calculated between its output and the ground-truth labels. In contrast, the teacher model is updated by taking an Exponential Moving Average (EMA) of the student model's weights during the training stage as follows.

$$\theta_t = \mu\theta_t + (1-\mu)\theta_s \tag{1}$$

where θ_t and θ_s are the parameters of the teacher model and the student model, and μ is a momentum coefficient. This process makes the teacher model a more robust and reliable source of pseudo-labels, as it averages out the rapid fluctuations of the student's training process. This approach allows the teacher model to provide a more stable and progressively refined target distribution throughout the training process. For unlabeled data, it is passed through different augmentations or perturbations and then fed into both the Teacher and student models separately. This process yields two sets of prediction probability maps, and a consistency loss is calculated between them to facilitate learning from the unlabeled data.

Building upon the teacher-student architecture, more recent semi-supervised approaches like DyCon (Assefa et al., 2025) introduces additional Uncertainty-aware Consistency Loss (UnCL) and the Focal Entropy-aware Contrastive Loss (FeCL). At a global scale, UnCL integrates voxel-wise uncertainty directly into the consistency loss via an entropy-driven dynamic weighting mechanism. While these general-propose semi-supervised methods have demonstrated further advancements on many segmentation benchmarks, their success is often measured on the segmentation of well-defined organs, causing them to overlook the critical topological structures required for more complex tasks like IA-Vessel segmentation. This issue is compounded because most existing models are evaluated on region overlap-based regularization. This metric is insensitive to geometric and topological abnormalities such as vessel adhesion and surface irregularities. Consequently, this often results in segmentation outcomes that are unsuitable for downstream applications due to subsequent mesh generation failures or flow field distortions.

2.3. Topological-Aware Regularization

Vascular networks are fundamentally tubular structures with complex topological properties, where connectivity and branching patterns represent their core anatomical features. However, existing semi-supervised segmentation methods primarily leverage unlabeled data through pixel-level or feature-level consistency, often neglecting this complex structural information. To address this challenge, we propose a plug-and-play topology-aware regularization loss, denoted as L_{Topo}, to enhance the model's awareness of structural integrity. The L_{Topo} is composed of a weighted sum of two complementary loss functions to optimize the vessel's topological structure from the perspectives of centerline matching with $\mathcal{L}_{clDice}$ and skeleton integrity $\mathcal{L}_{Skel}$ as follows:

$$\mathcal{L}_{\text{Topo}} = \lambda_1 \underbrace{\left(-\frac{2\sum_i \mathcal{T}(p_i)\mathcal{T}(g_i)}{\sum_i \mathcal{T}(p_i) + \sum_i \mathcal{T}(g_i)}\right)}_{\mathcal{L}_{\text{clDice}}} + \lambda_2 \underbrace{\left(-\frac{\sum_i p_i \mathcal{S}(g_i)}{\sum_i g_i}\right)}_{\mathcal{L}_{\text{Skel}}} \tag{2}$$

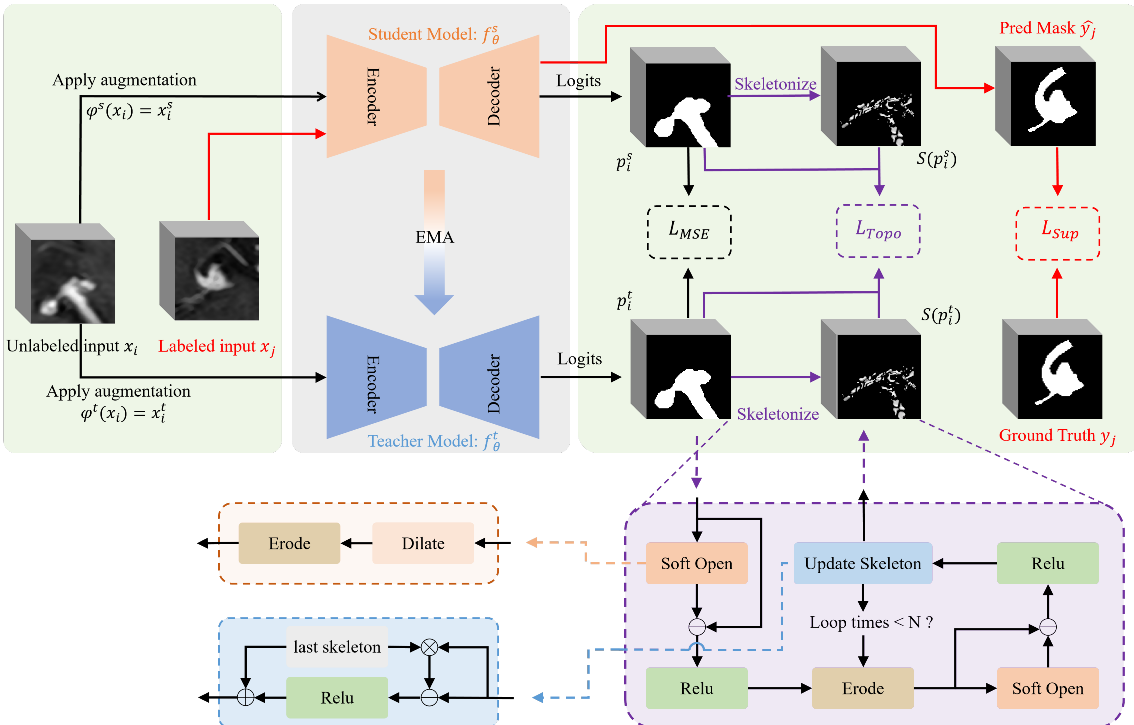

Figure 1: Overview of our proposed Topological-Aware Regularization framework for semi-supervised intracranial aneurysm vessel segmentation framework. For labeled data, a standard supervised segmentation loss ($\mathcal{L}_{\text{sup}}$) is applied to the student model's predictions. For unlabeled data, the teacher model's outputs are used as pseudo-labels to compute both the standard consistency loss ($\mathcal{L}_{\text{MSE}}$) and the topology-aware regularization loss ($\mathcal{L}_{\text{Topo}}$) that we propose. The figure also details our differentiable soft skeletonization algorithm, which approximates morphological erosion and dilation operations through a series of min-pooling and max-pooling operations. The $\mathcal{L}_{\text{Topo}}$ loss incorporates $\mathcal{L}_{\text{clDice}}$ and $\mathcal{L}_{\text{Skel}}$, which jointly penalize topological discrepancies between the student model's predictions and the teacher-generated pseudo-labels.

where p_i is the predicted probability at voxel i from the student model and g_i is the pseudo-label at voxel i from the teacher model. λ_1, λ_2 are weighting coefficients that balance the relative importance of the two component losses. $\mathcal{S}(\cdot)$ represents the soft-skeletonization function, which takes a probability map and outputs a map highlighting the central skeleton of the structure. $\mathcal{T}(\cdot)$ calculates the centerline probability map as follows.

$$\mathcal{T}(p_i) = \frac{\sum_i g_i \mathcal{S}(p_i)}{\sum_i \mathcal{S}(p_i)}, \quad \mathcal{T}(g_i) = \frac{\sum_i p_i \mathcal{S}(g_i)}{\sum_i \mathcal{S}(g_i)} \tag{3}$$

In semi-supervised framework, the binarized pseudo-labels g generated by the teacher model are treated as masks to supervise the student model's predictions p.

The original skeleton extraction algorithm is CPU-based (Kirchhoff et al., 2024). While its performance is sufficient in fully-supervised settings, this approach becomes computationally prohibitive for semi-supervised tasks where the teacher's pseudo-labels must be generated in real-time. To overcome this limitation, we adopted the differentiable soft-skeletonization as proposed in the clDice (Shit et al., 2021), to create a unified generation process for the skeletons required by both regularization losses. This algorithm efficiently approximates the vessel centerline skeleton on the GPU via a series of max-pooling and min-pooling operations, while maintaining full differentiability. Consequently, this unified design not only mitigates computational overhead but also ensures that both topological constraints operate on a consistent structural representation, thereby enhancing the model's capacity for learning both vessel connectivity and structural integrity.

2.4. Overall Training Procedure

The overall training objective of our proposed framework is to minimize the weighted sum of supervised segmentation loss $\mathcal{L}_{\text{sup}}$, unsupervised regularization loss of semi-supervised backbone $\mathcal{L}_{\text{unsup}}$ and our proposed topological-aware regularization $\mathcal{L}_{\text{Topo}}$ as follows.

$$\mathcal{L} = \mathcal{L}_{\text{sup}} + \mathcal{L}_{\text{unsup}} + \mathcal{L}_{\text{Topo}} \tag{4}$$

3. Experiments

3.1. Datasets

We conduct extensive experiments on the Intracranial Aneurysm Vessel Segmentation (IAVS) dataset (Xiao et al., 2025), which contains multi-center collection of 641 high-resolution 3D MRA images. In total, 587 IAs and their corresponding parent vessels were annotated and selected out to form patch volumes for training and validation of the sergmentation framework. The topological integrity of every vessel is guaranteed during the annotation procedure, providing a gold standard for validating the topology-preserving capabilities of segmentation models. We randomly partitioned the dataset into 357 cases for training, 99 cases for validation, and 66 cases for final testing evaluation. In our experiments, we use 5%, 10% and 20% of the training set (17, 35, and 71 cases) as labeled data, while the remaining cases served as unlabeled data, for which only the images were used during training. We compared our method against a series state-of-the-art semi-supervised segmentation methods. All methods were trained and evaluated under the identical labeled data configuration to ensure a fair comparison.

3.2. Implementation Details

All of our experiments are implemented in Python with PyTorch, using an NVIDIA A100 GPU. The backbone segmentation network for the specialist model is 3D U-NET (Çiçek et al., 2016). We use the SGD optimizer with an initial learning rate of 0.01, a weight decay of 1e-4 and a momentum of 0.9 to update the network parameters with the maximum iteration number set to 10000. To quantitatively evaluate the performance of all methods, we employed four standard metrics to assess segmentation accuracy from different perspectives. Dice Similarity Coefficient (Dice) is used to measure the volumetric overlap between the predicted segmentation and the ground truth. To quantify the discrepancy in volume, we use the Relative Absolute Volume Difference (RAVD). Furthermore, we evaluate the surface-to-surface accuracy using the Average Surface Distance (ASD) and the 95th percentile of the Hausdorff Distance (95HD). The ASD measures the average distance between the boundaries of the predicted and ground truth objects, while the 95HD provides a more robust measure of the maximum surface distance by excluding the top 5% of outlier distances. For both surface distance metrics, lower values indicate better performance.

3.3. Comparison Experiments

Table. 1 presents the performance of our method with comparison to other representative semi-supervised frameworks (Hung et al., 2018; Chen et al., 2021; Wu et al., 2021; Xu et al., 2023; Verma et al., 2022; Yu et al., 2019; Vu et al., 2019; Zhang et al., 2023; Wu et al., 2024) using different numbers of labeled images. An important observation from the results is the limited efficacy of several classic semi-supervised methods when applied to vessel segmentation. Notably, several methods such as ADV and CPS perform even worse than the Supervised Baseline trained with only the labeled data. This phenomenon highlights a critical challenge that methods that rely purely on region-based consistency are ill-suited for tasks where topological integrity is paramount. These conventional approaches enforce consistency at the pixel or patch level, which can be counterproductive for vessel networks and may penalize small valid gaps between different vessel segments. As a result, they may introduce misleading supervisory signals that corrupt the learning process.

In contrast, building upon the vanilla mean teacher framework (Tarvainen and Valpola, 2017), utilizing TAR consistently enhances the performance in all annotation scenarios. To further validate its impact on the leading edge of current research, we conduct experiments on DyCON (Assefa et al., 2025), a recently proposed state-of-the-art semi-supervised framework. As demonstrated in the table, the addition of TAR provides further performance increase and establishes a new state-of-the-art, showcasing its power as a plug-and-play module. This initial finding confirms that our topology-aware module provides a substantial and meaningful improvement to established semi-supervised methods. From the visualization of segmentation results in Figure. 2 We can observe that our proposed method generates more accurate predictions compared with other methods, which further demonstrates the effectiveness of our proposed method.

Table 1: Comparative experimental results between our proposed method and other semi-supervised segmentation methods on IAVS dataset with 5%, 10% and 20% annotation settings.

Method	**Annotation**	**Dice** [%]	**RAVD** [%]	**ASD**[voxel]	**95HD**[voxel]
Supervised Baseline	5%	66.03	39.83	1.79	37.70
ADV (BMVC'18)	5%	64.10	39.10	1.72	36.13
CPS (CVPR'21)	5%	64.37	45.97	1.63	28.70
RD (NeuriPS'21)	5%	65.70	38.48	1.69	32.82
ACMT (MedIA'23)	5%	64.72	47.31	1.74	36.83
ICT (NN'22)	5%	65.50	40.60	1.73	33.88
UAMT (MICCAI'19)	5%	66.15	39.91	1.67	3.54
EM (CVPR'19)	5%	65.02	39.92	1.71	33.96
UGMCL (AIIM'23)	5%	65.54	39.23	1.74	37.48
CML (ACMMM'24)	5%	63.85	39.96	1.87	15.77
MT (NeuriPS'17)	5%	66.14	39.32	1.76	34.80
MT + TAR	5%	**67.99**	**42.05**	**1.69**	**38.92**
DyCON (CVPR'25)	5%	67.40	40.26	1.56	20.85
Dycon + TAR	5%	**70.77**	**33.34**	**1.48**	**23.55**
Supervised Baseline	10%	66.64	41.02	3.10	16.70
ADV (BMVC'18)	10%	64.72	42.27	2.82	16.65
CPS (CVPR'21)	10%	65.96	40.36	3.28	17.83
RD (NeuriPS'21)	10%	66.44	36.38	2.93	17.17
ACMT (MedIA'23)	10%	67.13	38.52	3.20	16.51
ICT (NN'22)	10%	67.14	36.74	3.03	17.01
UAMT (MICCAI'19)	10%	67.57	37.83	2.85	17.53
EM (CVPR'19)	10%	67.62	39.69	3.32	16.38
UGMCL (AIIM'23)	10%	67.81	38.81	2.56	16.95
CML (ACMMM'24)	10%	68.99	17.94	2.03	15.70
MT (NeuriPS'17)	10%	67.10	40.36	3.12	16.75
MT + TAR	10%	**70.07**	**33.83**	**3.71**	**15.55**
DyCON (CVPR'25)	10%	68.23	38.68	2.18	16.32
Dycon + TAR	10%	**74.26**	**28.54**	**2.42**	**13.20**
Supervised Baseline	20%	72.72	34.72	1.54	32.99
ADV (BMVC'18)	20%	74.00	31.42	1.45	26.09
CPS (CVPR'21)	20%	74.85	31.63	1.51	33.22
RD (NeuriPS'21)	20%	74.65	30.10	1.38	32.23
ACMT (MedIA'23)	20%	73.73	30.90	1.54	31.88
ICT (NN'22)	20%	73.94	28.55	1.47	29.22
UAMT (MICCAI'19)	20%	74.21	30.98	1.46	31.57
EM (CVPR'19)	20%	73.18	34.99	1.42	27.72
UGMCL (AIIM'23)	20%	74.42	26.36	1.53	32.19
CML (ACMMM'24)	20%	74.14	30.03	1.36	32.66
MT (NeuriPS'17)	20%	74.68	28.08	1.45	31.49
MT + TAR	20%	**75.21**	**27.33**	**1.39**	**29.36**
DyCON (CVPR'25)	20%	74.20	27.69	1.32	16.83
Dycon + TAR	20%	**76.81**	**22.50**	**1.22**	**22.09**
Supervised Upperbound	100%	79.00	24.07	2.47	11.95

Table 2: Ablation experiments of different components of topological-aware regularization on Dycon semi-supervised segmentation framework using 10% labeled data.

Method	**Dice** [%]	**RAVD** [%]	**ASD**[voxel]	**95HD**[voxel]	**Iteration time**[s]
Baseline w/o TAR	68.23	38.68	2.18	16.32	0.37
BettiMatching	61.47	46.07	2.25	19.06	9.20
clDice	69.39	34.95	**2.11**	15.84	0.51
Skel	71.75	32.34	2.25	14.03	1.09
cl-Skel	73.73	28.61	2.57	**12.70**	0.40
clDice+BettiMatching	65.00	49.61	2.37	17.50	9.30
clDice+Skel	72.77	31.45	2.21	13.88	1.26
clDice+cl-Skel (Ours)	**74.26**	**28.54**	2.42	13.20	**0.79**

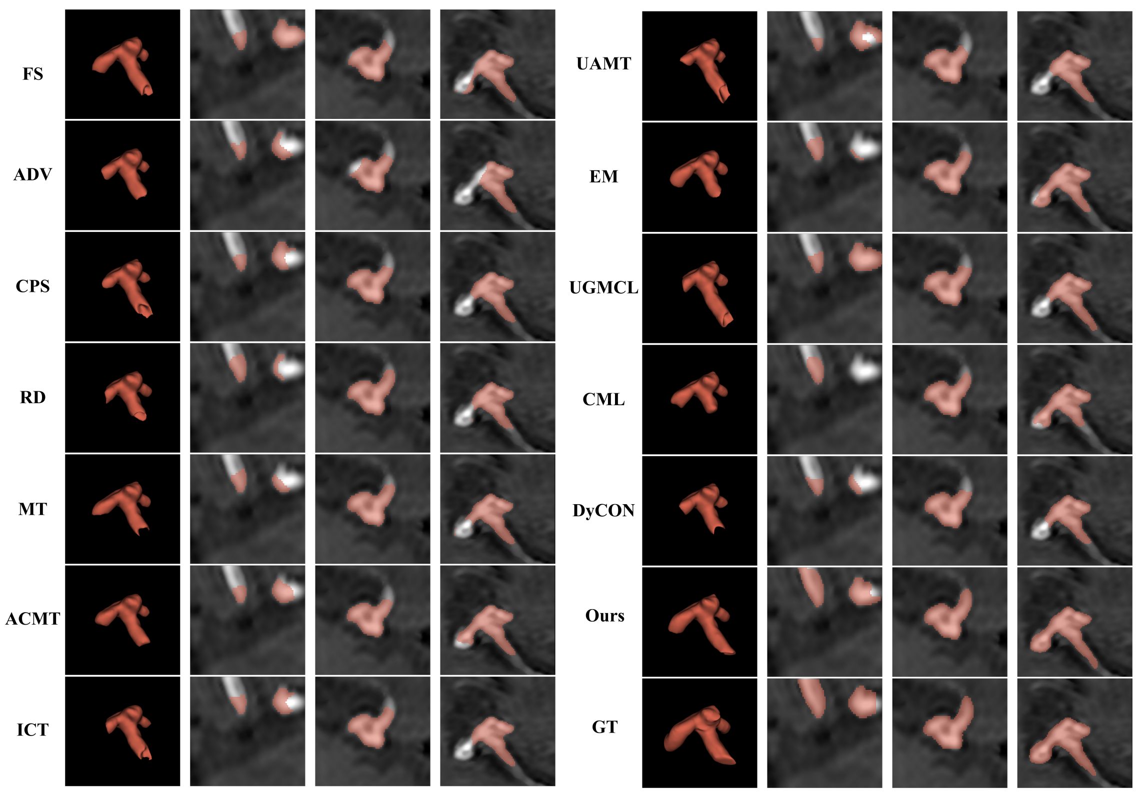

Figure 2: Visual comparison of the intracranial aneurysm vessel results of our proposed method with comparison to fully supervised baseline and other state-of-the-art semi-supervised methods.

3.4. Ablation analysis

To validate the contribution of each component within TAR and demonstrate the critical impact of our efficient implementation, we conducted a thorough ablation analysis with comparison to other topological-aware regularization strategies (Stucki et al., 2024) in Table. 2. The results validate that our selection of utilizing both the clDice loss and the Skeleton loss (Skel) achieves the best performance on most metrics including Dice and RAVD. However, the original CPU-based skeletonization (Skel) introduces a severe computational bottleneck, more than doubling the iteration time. The proposed regularization, leveraging the differentiable soft-skeleton algorithm from clDice (cl-Skel), concurrently improves segmentation accuracy and computational efficiency.

4. Conclusion and Discussion

In this study, we proposed and validated the core hypothesis that explicitly incorporating topological structure priors is crucial for semi-supervised vessel segmentation, with experimental results providing strong support. Our Topological-Aware Regularization (TAR), a plug-and-play component, demonstrated significant performance improvements when integrated with different semi-supervised frameworks, proving its efficiency in guiding the model to learn the intrinsic structure of vascular networks. Our findings reveal a key limitation of current semi-supervised learning methods. As shown in the experiments, several classic methods relying on region-based consistency perform even worse than the supervised baseline trained solely on labeled data when handling complex vascular structures. This indicates that for tasks with strong structural priors, blindly enforcing pixel/feature-level consistency can generate misleading supervisory signals that impair learning. Our work advocates for a paradigm shift from generic, task-agnostic consistency regularization to semantic constraints integrated with specific anatomical priors.

Although our proposed method achieves significant improvements, it still has limitations. Firstly, the effectiveness of TAR remains partially dependent on the quality of pseudo-labels generated by the teacher model. When labeled data is scarce, incorrect topological structures might be treated as supervisory signals, thus affecting the student model's learning. Besides, incorporating more complex topological descriptors into the consistency learning framework could enable more comprehensive structural preservation (Lux et al., 2024; Stucki et al., 2024). Finally, we will explore applying the TAR framework to other medical image segmentation tasks involving tubular or network-like structures to validate its generality and effectiveness (Yao et al., 2024; Tan et al., 2022; Sun et al., 2023; Xue et al., 2020). As a preliminary exploration of integrating topological structure priors into semi-supervised segmentation, our work lays a foundation and provides a new perspective for addressing complex anatomical segmentation tasks with limited labeled data.

References

Michela Antonelli, Annika Reinke, Spyridon Bakas, Keyvan Farahani, Annette Kopp-Schneider, Bennett A Landman, Geert Litjens, Bjoern Menze, Olaf Ronneberger, Ronald M Summers, et al. The medical segmentation decathlon. *Nature communications*, 13(1):4128, 2022.

Maregu Assefa, Muzammal Naseer, Iyyakutti Iyappan Ganapathi, Syed Sadaf Ali, Mohamed L Seghier, and Naoufel Werghi. Dycon: Dynamic uncertainty-aware consistency and contrastive learning for semi-supervised medical image segmentation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 30850–30860, 2025.

Juan R Cebral, Marcelo Adrián Castro, Sunil Appanaboyina, Christopher M Putman, Daniel Millan, and Alejandro F Frangi. Efficient pipeline for image-based patient-specific analysis of cerebral aneurysm hemodynamics: technique and sensitivity. *IEEE transactions on medical imaging*, 24(4):457–467, 2005.

Xiaokang Chen, Yuhui Yuan, Gang Zeng, and Jingdong Wang. Semi-supervised semantic segmentation with cross pseudo supervision. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2613–2622, 2021.

Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.

Nima Etminan and Gabriel J Rinkel. Unruptured intracranial aneurysms: development, rupture and preventive management. *Nature Reviews Neurology*, 12(12):699–713, 2016.

Wei Chih Hung, Yi Hsuan Tsai, Yan Ting Liou, Yen Yu Lin, and Ming Hsuan Yang. Adversarial learning for semi-supervised semantic segmentation. In *29th British Machine Vision Conference, BMVC 2018*, 2018.

Rushi Jiao, Yichi Zhang, Le Ding, Bingsen Xue, Jicong Zhang, Rong Cai, and Cheng Jin. Learning with limited annotations: a survey on deep semi-supervised learning for medical image segmentation. *Computers in Biology and Medicine*, page 107840, 2023.

Yannick Kirchhoff, Maximilian R Rokuss, Saikat Roy, Balint Kovacs, Constantin Ulrich, Tassilo Wald, Maximilian Zenk, Philipp Vollmuth, Jens Kleesiek, Fabian Isensee, et al. Skeleton recall loss for connectivity conserving and resource efficient segmentation of thin tubular structures. In *European Conference on Computer Vision*, pages 218–234. Springer, 2024.

Xigui Li, Yuanye Zhou, Feiyang Xiao, Xin Guo, Yichi Zhang, Chen Jiang, Jianchao Ge, Xiansheng Wang, Qimeng Wang, Taiwei Zhang, et al. Aneumo: A large-scale comprehensive synthetic dataset of aneurysm hemodynamics. *arXiv preprint arXiv:2501.09980*, 2025.

Yuexiang Li, Jiawei Chen, Xinpeng Xie, Kai Ma, and Yefeng Zheng. Self-loop uncertainty: A novel pseudo-label for semi-supervised medical image segmentation. In *International conference on medical image computing and computer-assisted intervention*, pages 614–623. Springer, 2020.

Liyun Lu, Mengxiao Yin, Liyao Fu, and Feng Yang. Uncertainty-aware pseudo-label and consistency for semi-supervised medical image segmentation. *Biomedical Signal Processing and Control*, 79:104203, 2023.

Xiangde Luo, Guotai Wang, Wenjun Liao, Jieneng Chen, Tao Song, Yinan Chen, Shichuan Zhang, Dimitris N Metaxas, and Shaoting Zhang. Semi-supervised medical image segmentation via uncertainty rectified pyramid consistency. *Medical Image Analysis*, 80: 102517, 2022.

Laurin Lux, Alexander H Berger, Alexander Weers, Nico Stucki, Daniel Rueckert, Ulrich Bauer, and Johannes C Paetzold. Topograph: An efficient graph-based framework for strictly topology preserving image segmentation. *arXiv preprint arXiv:2411.03228*, 2024.

Jun Ma, Yao Zhang, Song Gu, Cheng Zhu, Cheng Ge, Yichi Zhang, Xingle An, Congcong Wang, Qiyuan Wang, Xin Liu, Shucheng Cao, Qi Zhang, Shangqing Liu, Yunpeng Wang, Yuhui Li, Jian He, and Xiaoping Yang. Abdomenct-1k: Is abdominal organ segmentation a solved problem? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6695–6714, 2022.

Paul D Morris, Andrew Narracott, Hendrik von Tengg-Kobligk, Daniel Alejandro Silva Soto, Sarah Hsiao, Angela Lungu, Paul Evans, Neil W Bressloff, Patricia V Lawford, D Rodney Hose, et al. Computational fluid dynamics modelling in cardiovascular medicine. *Heart*, 102(1):18–28, 2016.

Tatsat R Patel, Aakash Patel, Sricharan S Veeturi, Munjal Shah, Muhammad Waqas, Andre Monteiro, Ammad A Baig, Nandor Pinter, Elad I Levy, Adnan H Siddiqui, et al. Evaluating a 3d deep learning pipeline for cerebral vessel and intracranial aneurysm segmentation from computed tomography angiography–digital subtraction angiography image pairs. *Neurosurgical Focus*, 54(6):E13, 2023.

Laurent Pierot, Christophe Portefaix, Christine Rodriguez-Régent, Sophie Gallas, Jean-François Meder, and Catherine Oppenheim. Role of mra in the detection of intracranial aneurysm in the acute phase of subarachnoid hemorrhage. *Journal of Neuroradiology*, 40(3):204–210, 2013.

Chongyu Qu, Tiezheng Zhang, Hualin Qiao, Yucheng Tang, Alan L Yuille, Zongwei Zhou, et al. Abdomenatlas-8k: Annotating 8,000 ct volumes for multi-organ segmentation in three weeks. *Advances in Neural Information Processing Systems*, 36:36620–36636, 2023.

Wouter I Schievink. Intracranial aneurysms. *New England Journal of Medicine*, 336(1): 28–40, 1997.

Yuyan Shi, Yichi Zhang, and Shasha Wang. Competitive ensembling teacher-student framework for semi-supervised left atrium mri segmentation. In *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 2231–2234. IEEE, 2023.

Yuyan Shi, Jialu Ma, Jin Yang, Shasha Wang, and Yichi Zhang. Beyond pixel-wise supervision for medical image segmentation: From traditional models to foundation models. *arXiv preprint arXiv:2404.13239*, 2024.

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.

Nico Stucki, Vincent Bürgin, Johannes C Paetzold, and Ulrich Bauer. Efficient betti matching enables topology-aware 3d segmentation via persistent homology. *arXiv preprint arXiv:2407.04683*, 2024.

Shenghuan Sun, Greg Goldgof, Atul Butte, and Ahmed M. Alaa. Aligning synthetic medical images with clinical knowledge using human feedback. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, pages 13408–13428. Curran Associates, Inc., 2023.

Nima Tajbakhsh, Laura Jeyaseelan, Qian Li, Jeffrey N Chiang, Zhihao Wu, and Xiaowei Ding. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. *Medical Image Analysis*, 63:101693, 2020.

Wenjun Tan, Pan Liu, Xiaoshuo Li, Shaoxun Xu, Yufei Chen, and Jinzhu Yang. Segmentation of lung airways based on deep learning methods. *IET Image Processing*, 16(5): 1444–1456, 2022.

Antti Tarvainen and Harri Valpola. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. volume 30, 2017.

Bethany H Thompson, Gaetano Di Caterina, and Jeremy P Voisey. Pseudo-label refinement using superpixels for semi-supervised brain tumour segmentation. In *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, pages 1–5. IEEE, 2022.

Vikas Verma, Kenji Kawaguchi, Alex Lamb, Juho Kannala, Arno Solin, Yoshua Bengio, and David Lopez-Paz. Interpolation consistency training for semi-supervised learning. *Neural Networks*, 145:90–106, 2022.

Tuan-Hung Vu, Himalaya Jain, Maxime Bucher, Matthieu Cord, and Patrick Pérez. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2517–2526, 2019.

Jing Wang, Fengtao Liu, Jingjie Ge, Yimin Sun, Yilin Tang, Xiaoniu Liang, Yixin Zhao, Wenbo Yu, Jianjun Wu, Chensen Lin, et al. Nasal septum deviation correlates with asymmetry in parkinson's disease. *medRxiv*, pages 2025–02, 2025.

Lijun Wu, Juntao Li, Yue Wang, Qi Meng, Tao Qin, Wei Chen, Min Zhang, Tie-Yan Liu, et al. R-drop: Regularized dropout for neural networks. *Advances in neural information processing systems*, 34:10890–10905, 2021.

Song Wu, Xiaoyu Wei, Xinyue Chen, Yazhou Ren, Jing He, and Xiaorong Pu. Cross-view mutual learning for semi-supervised medical image segmentation. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pages 9253–9261, 2024.

Feiyang Xiao, Yichi Zhang, Xigui Li, Yuanye Zhou, Chen Jiang, Xin Guo, Limei Han, Yuxin Li, Fengping Zhu, and Yuan Cheng. Rethinking intracranial aneurysm vessel segmentation: A perspective from computational fluid dynamics applications. *arXiv preprint arXiv:2512.01319*, 2025.

Zhe Xu, Yixin Wang, Donghuan Lu, Xiangde Luo, Jiangpeng Yan, Yefeng Zheng, and Raymond Kai-yu Tong. Ambiguity-selective consistency regularization for mean-teacher semi-supervised medical image segmentation. *Medical Image Analysis*, 88:102880, 2023.

Yuan Xue, Nan Du, Anne Mottram, Martin Seneviratne, and Andrew M Dai. Learning to select best forecast tasks for clinical outcome prediction. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 15031–15041. Curran Associates, Inc., 2020.

Huifeng Yao, Xiaowei Hu, and Xiaomeng Li. Enhancing pseudo label quality for semi-supervised domain-generalized medical image segmentation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, pages 3099–3107, 2022.

Linlin Yao, Dongdong Chen, Xiangyu Zhao, Manman Fei, Zhiyun Song, Zhong Xue, Yiqiang Zhan, Bin Song, Feng Shi, Qian Wang, et al. Aaseg: Artery-aware global-to-local framework for aneurysm segmentation in head and neck cta images. *IEEE Transactions on Medical Imaging*, 2024.

Lequan Yu, Shujun Wang, Xiaomeng Li, Chi-Wing Fu, and Pheng-Ann Heng. Uncertainty-aware self-ensembling model for semi-supervised 3d left atrium segmentation. In *International conference on medical image computing and computer-assisted intervention*, pages 605–613. Springer, 2019.

Yichi Zhang, Rushi Jiao, Qingcheng Liao, Dongyang Li, and Jicong Zhang. Uncertainty-guided mutual consistency learning for semi-supervised medical image segmentation. *Artificial Intelligence in Medicine*, 138:102476, 2023.

Yichi Zhang, Zhenrong Shen, and Rushi Jiao. Segment anything model for medical image segmentation: Current applications and future directions. *Computers in Biology and Medicine*, 171:108238, 2024a.

Yichi Zhang, Jin Yang, Yuchen Liu, Yuan Cheng, and Yuan Qi. Semisam: Enhancing semi-supervised medical image segmentation via sam-assisted consistency regularization. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 3982–3986. IEEE, 2024b.

Yichi Zhang, Bohao Lv, Le Xue, Wenbo Zhang, Yuchen Liu, Yu Fu, Yuan Cheng, and Yuan Qi. Semisam+: Rethinking semi-supervised medical image segmentation in the era of foundation models. *Medical Image Analysis*, 2025.