VesselSim: learning 3D blood vessel segmentation without expert annotations

Erin Rainville[1⋆], Melissa Ananian[1], Tristan Mirolla[1], Hassan Rivaz[2], and Yiming Xiao[1]

[1] Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada
[2] Department of Computer Science and Software Engineering, Concordia University, Montreal, Canada
e_ainvil@live.concordia.ca

Abstract. Blood vessel segmentation is a core task in medical image analysis for the care of vascular diseases and surgical planning, yet the challenges of providing expert vascular annotations pose a major obstacle for the progress of related deep learning techniques. To address this, we propose `VesselSim`, a two-stage framework for universal 3D blood vessel segmentation that eliminates the need for real annotated data during training. First, we introduce a stochastic, geometry-driven vascular simulation framework that models recursive branching, curvature-controlled growth, and collision-aware topology, followed by domain-randomized intensity synthesis to generate 16,500 anatomically plausible 3D angiographic volumes. Second, a 3D U-Net is trained solely on this synthetic data. To bridge the domain gap from synthetic to real images at inference time, we introduce a test-time adaptation strategy via a self-supervised mask reconstruction decoder, enabling adaptation to unseen clinical scans without prior domain knowledge. We evaluate `VesselSim` in a zero-shot setting on multiple real-world datasets spanning MR and CT across different anatomical regions, including the brain and kidneys. Despite being trained exclusively on synthetic data, `VesselSim` achieves performance competitive with state-of-the-art vascular segmentation foundation models. These findings suggest that learning vessel geometry from synthetic tubular structures is effective for robust cross-domain generalization, substantially reducing the reliance on acquired medical imaging data and more importantly, expert annotations.

Keywords: Vascular segmentation · Synthetic data · Test-time training.

1 Introduction

Blood vessel segmentation from 3D medical scans plays a critical role in various clinical applications, including diagnosing vascular disorders [16], planning surgical interventions [3], and studying disease-related perfusion abnormality as

⋆ Corresponding author.

biomarkers [12]. With delicate geometry, soft contrast, complex anatomical context, and widespread spatial distribution, manual labeling of vasculature is often much more challenging and time-consuming than other body organs. Largely due to this, publicly labeled 3D vascular datasets are scarce, greatly hindering the development of deep-learning (DL)-based vascular segmentation methods.

To address this, recent advances in data-efficient 3D vessel segmentation focus on reducing annotation demand and improving domain generalization. Few-shot learning approaches tailored to vascular structures have shown promise. VesselShot [1] explores few-shot cerebrovascular segmentation by leveraging knowledge from few labeled support samples. FSVS-Net [26] introduces a semi- supervised few-shot framework that uses feature distillation and bidirectional weighted fusion to propagate sparse annotated slices to unlabeled ones for hepatic, pulmonary, and renal vessel segmentation. Later, biomedical foundation models have also advanced data efficiency for the task. Notably, VesselFM [23] is a universal foundation model specifically trained for 3D blood vessel segmentation using a combination of domain-randomized synthetic images, flow-matching generated samples, and real annotated data. It enables robust zero- and few-shot performance across diverse scans. Besides algorithm development, to further mitigate the dependency on expert annotations, a few rule- and DL-based approaches for vessel simulation [8, 7, 11, 13, 21] have also been proposed, but they are often domain-specific (with reliance on training data) or lack typical curvy geometries found in vasculatures, potentially limiting their generalization capacity. In addition to training samples, domain shift poses another challenge for the robustness of vessel segmentation algorithms. More recently, test-time adaptation (TTA) methods [27] that use unlabeled data to efficiently align model features or regularize outputs at inference time have shown great power in medical image segmentation. Yet, despite early investigations with 2D scans [10, 28], *very few have explored TTA in the realm of 3D blood vessel segmentation.*

As existing data-efficient 3D vessel segmentation methods still rely on real vascular annotations and lack robustness to domain shift, we propose `VesselSim`, a new blood vessel segmentation DL model trained on data synthesized by a probabilistic rules-based simulation for more natural looking geometries, without manual segmentation ground truths. Our two-stage framework overcomes the aforementioned limitations by combining topology-constrained synthetic pretraining with ground-truth-free test-time adaptation, enabling structurally consistent segmentation and improved generalization from synthetic to real clinical volumes. *Our major contributions are three-fold*: **First**, we proposed a new stochastic technique to simulate artificial 3D angiography scans at different scales for DL model training, and release the code and dataset to the community at *https://healthx-lab.github.io/VesselSim*. **Second**, we investigate a novel test-time adaptation technique based on mask reconstruction to efficiently bridge domain shifts between synthetic and real clinical data without needing voxel-wise annotations. **Lastly**, we compare the proposed method against state-of-the-art foundation models across multiple 3D vascular segmentation tasks.

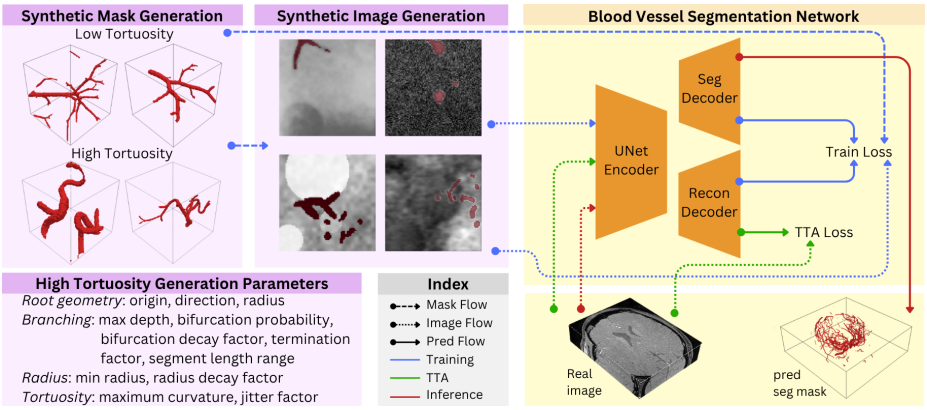

Fig. 1. `VesselSim` framework. Left: Synthetic vascular masks are generated with controllable geometric parameters. Masks are converted into angiography-like volumes via domain-randomization. Right: A 3D U-Net trained on the synthetic data. At inference, the self-supervised reconstruction decoder enables test-time adaptation to real scans.

2 Methods and Materials

The overall pipeline of the proposed angiography simulation method and deep learning framework with test-time adaptation for blood vessel segmentation are summarized in Fig. 1.

2.1 Artificial vascular image generation

Existing vascular simulation methods vary in realism and data dependence. Classical rule-based approaches [13, 9] rely on recursive branching or cost-minimization principles with the possibility of generating large, complex vascular networks, but they typically result in stick-like or piecewise-linear segments that insufficiently represent the tortuous/sinuous geometry of certain vasculature (e.g., in the brain). In contrast, deep generative models, such as VesselGPT [8] and VesselVAE [7] achieve more realistic cerebrovascular morphologies, but require large annotated clinical datasets and are restricted to short local vessel fragments.

To enable multi-scale, data-independent vascular synthesis, we propose a procedural stochastic branching framework. Extending from [13], we incorporate a biased random walk into a stochastic recursive branching system to control segment trajectories, enabling curvature and directional persistence. Segment length, rotation axis, and rotation magnitude are sampled from defined probabilistic ranges informed by the literature [24, 15], and growth proceeds in a *breadth-first manner* within a $160 \times 160 \times 160$ voxel volume to ensure uniform spatial density. Collision avoidance is enforced via multi-attempt path validation, terminating branches when no feasible path exists. Additionally, branching probability decays with depth while radii decrease following physiological scaling consistent with Murray's law [17]. *To generate the final angiography volume*

(see Fig. 1), we extract up to 50 random $96 \times 96 \times 96$-voxel sub-volumes from each main one that pass an automatic quality control based on criteria of flow continuity, vessel integrity (no small "floating islands"), and a 5% minimum vessel occupancy threshold. Then, we augment random 3D geometric shapes as the background of the vessel patches and add Gaussian noise to simulate MRI and CT angiographies, following VesselFM's domain randomization scheme [23]. Notably, we intentionally added ellipsoidal shell shapes to simulate skulls to improve performance. For the full training dataset, we combine 5000 low-tortuosity vessel samples with VascuSynth [13], 5000 high-tortuosity ones with our method, 2500 skull samples, and 2500 background samples to reduce false positive rates. An additional 500 low-tortuosity, 500 high-tortuosity, 250 skull-injected and 250 background samples are separately generated as the validation set to prevent data leakage between training and validation sets. We use real clinical vascular angiographies as testing data, with details described in Section 2.4.

2.2 Learning with synthetic vascular data

`VesselSim` is based on a 3D UNet architecture [19] trained on synthetic data from our vessel generation framework in an end-to-end manner. To encourage accurate segmentation of thin, branching vascular structures, we employ a composite objective function. The primary segmentation loss consists of Cross-Entropy ($\mathcal{L}_{CE}$) and Dice loss ($\mathcal{L}_{Dice}$), balancing voxel-wise classification with volumetric overlap. To further penalize topological discontinuities and centerline fragmentation, we integrate a centerline boundary Dice loss ($\mathcal{L}_{cbDice}$) [20]. This term leverages soft-skeletonization to enforce alignment between predicted and ground-truth vessel centerlines and boundaries, providing an explicit inductive bias toward tubular geometry and connectivity.

To facilitate domain adaptation from synthetic to real images, we extend the segmentation network with an auxiliary self-supervised mask reconstruction task. We attach a reconstruction decoder to the shared encoder bottleneck to perform in-painting. During training, input volumes are corrupted by randomly masking multiple cubical regions (ranging from $2\times2\times2$ to $16\times16\times16$ voxels) [18, 5]. The network is jointly optimized for segmentation and reconstruction using an L_1 reconstruction loss ($\mathcal{L}_{rec}$). The reconstruction loss is computed only over the masked voxels, encouraging the network to infer missing vascular structures from surrounding context rather than learning an identity mapping. This auxiliary task provides a self-supervised objective that can later be leveraged for test-time adaptation. The total training objective is:

$$\mathcal{L}_{total} = \alpha\mathcal{L}_{CE} + \beta\mathcal{L}_{Dice} + \gamma\mathcal{L}_{cbDice} + \lambda\mathcal{L}_{rec} \tag{1}$$

where weights are empirically set to $\alpha = 0.1, \beta = 0.9, \gamma = 0.2$, and $\lambda = 0.1$. The model is trained exclusively on the generated dataset of 16,500 volumetric patches, using a single NVIDIA H100 GPU with a batch size of 16. Optimization is performed with an initial learning rate of 10^{-4}, followed by linear decay to 10^{-5} until convergence. All input image intensities are normalized to [0,1].

2.3 Test-time adaptation for vascular segmentation

We leverage the reconstruction decoder as a self-supervised test-time adaptation mechanism to mitigate domain shift between synthetic training data and real clinical images, which can differ in contrast, noise characteristics, intensity scaling, acquisition artifacts, and anatomical features. The reconstruction objective provides an unsupervised alignment signal that encourages the encoder to adjust its internal representations to the appearance characteristics of each target volume, without requiring any expert annotations.

TTA is performed independently for each unseen test volume. At inference, we randomly sample 24 patches ($96 \times 96 \times 96$ voxels) from the target volume, apply the masking strategy from training, and minimize $\mathcal{L}_{rec}$, without ever using segmentation labels. Optimization runs for a single epoch using the Adam optimizer (learning rate=10^{-4}). Crucially, only encoder instance normalization parameters are updated while all other weights remain frozen, efficiently recalibrating the activation distributions to the target domain while preserving the learned structural vascular priors. The final segmentation is then obtained by forwarding the uncorrupted test volume through the adapted segmentation model. As a final refinement step, small disconnected components below a predefined volume threshold are removed to suppress spurious non-vascular predictions.

2.4 Real vascular data and comparison implementation

To evaluate generalization beyond synthetic data, we perform inference on two publicly available medical image repositories, *HiP-CT* and *TopCoW* (3 distinct datasets in total), thus allowing us to assess robustness across organ systems (kidney and brain) and across imaging contrasts (HiP-CT, CTA and MRA).

The ***HiP-CT*** database [14] consists of three ultra–high-resolution 3D human kidney volumes acquired using Hierarchical Phase-Contrast Tomography. The original volumes have an average dimension of $1465 \times 1494 \times 1732$ voxels with an isotropic resolution from $5um$ to $50um$. Due to their large size, volumes are partitioned into $500 \times 500 \times 500$ voxel sub-volumes for computational feasibility, which are used consistently across all baseline models and experiments.

The ***TopCoW*** dataset [25] includes 125 computed tomography angiographies (CTA) and 125 magnetic resonance angiographies (MRA) of the human brains. These modalities differ substantially in contrast mechanisms and the image characteristics. While the CTA volumes have an average resolution of $0.45 \times 0.45 \times 0.7\ mm^3$, the MRA volumes average $0.3 \times 0.3 \times 0.6\ mm^3$. All scans are resampled to an common resolution of $0.4 \times 0.4 \times 0.4\, mm^3$. For ***TopCoW CTA***, manual annotations are provided for the Circle of Willis. Volumes are cropped to the bounding box of this region to avoid penalizing predictions outside the territory. For ***TopCoW MRA***, we use the VesselVerse extension [6], which expands the TopCoW MRA annotations to the full cerebrovasculature.

To contextualize the performance of `VesselSim` on these datasets, we compare against three recent state-of-the-art (SOTA) medical image segmentation

foundation models evaluated on the same test sets. UniverSeg [2] is a general-purpose prompt-based medical segmentation model whose capability for vascular segmentation has been previously demonstrated. It is trained on 53 heterogeneous medical segmentation datasets and performs task adaptation through a support-set prompting mechanism. For each target dataset, we provide four randomly selected volumes as a support set to condition the model on the segmentation task. We additionally evaluate SAM-Med3D [22], a 3D extension of the Segment Anything Model tailored for volumetric medical imaging. SAM-Med3D leverages large-scale pretraining and prompt-based mask generation to enable adaptable segmentation across anatomical structures. VesselFM [23] is a foundation model specialized for vascular segmentation, pretrained across 19 vessel datasets and applied here in a zero-shot manner without additional adaptation. Notably, both TopCoW and HiP-CT are included in the VesselFM training corpus, potentially conferring an inherent advantage in this evaluation. In contrast to all foundation baselines, `VesselSim` is trained exclusively on synthetic vascular data and has no exposure to real-world clinical images during training.

As the foundation models are pretrained on large-scale real-world datasets, we additionally evaluate whether `VesselSim` can benefit from limited exposure to real target-domain data. For each dataset, we perform a small-sample model finetuning using the same four support volumes used for UniverSeg. From each volume, we randomly sample 24 ($96 \times 96 \times 96$ voxel) patches, resulting in 96 training patches per dataset, then finetune for a single epoch using a learning rate of 10^{-5}, updating all network weights. This setting simulates a realistic low-annotation regime, where only a small number of labeled scans are available.

3 Results

Table 1 summarizes the comparisons between foundation model baselines and `VesselSim` variants. Additionally, Fig. 2 showcases some qualitative comparisons between our method and baselines. Performance is evaluated across all datasets using Dice for volumetric accuracy and clDice for topological consistency of tubular structures. Overall, the zero-shot `VesselSim` trained solely on our complete synthetic dataset achieves competitive performance relative to foundation models trained on substantially larger real clinical datasets. Notably, it generalizes well across both neurovascular (CTA/MRA) and renal CT domains, outperforming or matching the baselines. This demonstrates that synthetic-only training can transfer effectively to real clinical data across organs and imaging modalities.

HiP-CT represents an extreme domain shift due to its ultra-high-resolution imaging characteristics. In this setting, UniverSeg and SAM-Med3D fail to detect vessels reliably, resulting in near-zero Dice and clDice scores. VesselFM was pretrained on HiP-CT so it performs well. In contrast, `VesselSim` achieves competitive performance, demonstrating that geometry-driven synthetic training enables robust generalization even under severe appearance shifts.

In addition, in our ablation studies (see Table 1), we compare different `VesselSim` variants and reveal the contribution of different key training com-

Table 1. Segmentation performance across datasets. `VesselSim` is our final model (loss *CE + Dice + cbDice + Reconstruction TTA*). The best results for each dataset (Dice %, clDice %) are in bold, second best results are underlined. The last row denotes the small-sample finetuning upper bound of `VesselSim` and is excluded from highlighting.

Models	TopCoW CTA		TopCoW MRA		HiP-CT	
	Dice	clDice	Dice	clDice	Dice	clDice
VesselFM	**52.9±9.5**	58.9±10.2	48.3±8.1	46.9±9.1	35.1±24.5	25.6±19.5
SAM-Med3D	6.6±4.4	9.1±6.0	3.1±2.6	2.6±2.2	8.0±21.6	3.1±8.9
UniverSeg	19.7±8.6	29.4±12.3	48.2±5.1	47.0±6.1	4.7±21.1	0.0±0.0
U-Net (Base)	42.7±12.2	60.6±12.7	70.1±3.9	73.1±5.1	34.3±23.0	23.0±42.7
Base + cbDice	46.8±12.9	62.9±13.6	**72.6±3.5**	75.7±4.4	31.1±23.7	41.2±25.1
Base + recon	43.1±14.1	60.6±13.3	71.9±3.9	74.6±5.2	**36.0±22.7**	41.3±23.1
VesselSim	48.7±11.9	**64.0±13.0**	71.8±3.3	**76.2±4.1**	31.6±21.5	**42.3±23.8**
Finetuned	59.5±7.7	70.4±7.3	77.3±3.3	83.2±3.6	46.8±33.2	50.7±27.8

ponents. Starting from the 3D U-Net base model (*CE + Dice*), incorporating *cbDice* primarily improves topological consistency, which is reflected in systematic gains in *clDice* across modalities. In contrast, adding the reconstruction auxiliary task yields larger improvements in cross-domain CT performance, suggesting enhanced robustness to appearance shifts. `VesselSim`, with *cbDice* loss and reconstruction-based TTA, produces the strongest overall model, consistently improving structural preservation while maintaining or improving volumetric overlap. These complementary effects indicate that *cbDice* promotes geometric continuity of thin tubular structures, whereas reconstruction-based adaptation mitigates modality-specific intensity differences. Finally, small-sample finetuning can further boost performance across all datasets, particularly in the cross-organ CT setting. This indicates that synthetic pretraining provides a strong prior that can be effectively refined with minimal real supervision.

4 Discussion

The proposed `VesselSim` relies on training from 15,000 synthetic vascular 3D patches without any expert annotation, in contrast to baseline foundation models that are trained on substantially larger real-data corpora (e.g., VesselFM trained on 115,000 real angiography patches and extensive domain-randomized data simulated from expert annotations). Yet, `VesselSim` matches or surpasses VesselFM, SAM-Med3D, and UniverSeg on multiple datasets. A few design choices contribute to the performance and training efficiency. **First**, the proposed enhanced stochastic vessel generation algorithm in addition to domain randomization helps produce more diverse, natural looking vascular patches, which are beneficial for learning strong priors to compensate for the absence of large-scale real data supervision. Also, due to the challenges in manual annotation of blood

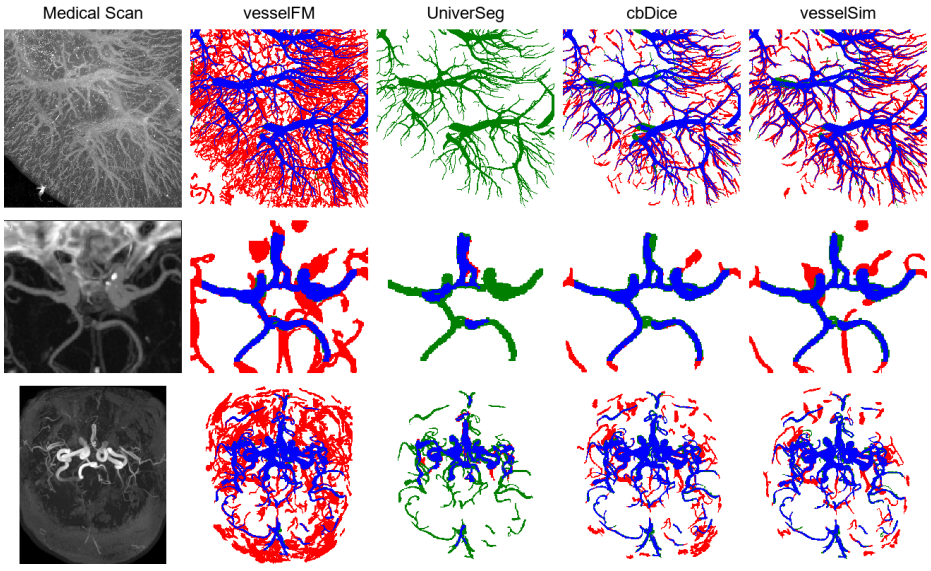

Fig. 2. Maximum intensity projections of the input scan and different model predictions. From top to bottom: HiP-CT, TopCoW CTA, TopCoW MRA. For the predictions, true positives are shown in blue, false positives in red, and false negatives in green. SAM-Med3D is omitted due to low segmentation performance across datasets.

vessels, expert labels can often miss true targets while synthetic data offers more confident ground truths. *Notably, with the additional skull simulation, the mean Dice score was improved from 0.61 to 0.72 before TTA for the TopCoW MRA dataset. Similarly, omitting high-tortuosity vessels during* `VesselSim` *training triggered a severe performance degradation. For TopCoW CTA, which consists of the tortuous Circle of Willis anatomy, the Dice and clDice scores fell by 26.8% and 28.4% respectively.* **Second**, with the addition of the mask reconstruction auxiliary task and *cbDice* loss, training emphasizes on learning tubular structural features. Notably, improvement in *clDice* across modalities further supports the hypothesis that the model learns geometry-driven representations, considering VesselFM does not include additional losses designed for tubular features. **Lastly**, by including the auxiliary task intended for TTA during main model training, it helps guide domain adaptation for segmentation [4]. We focused on adapting the instance normalization layers of the 3D U-Net rather than the full encoder to help mitigate "weight drift" under small sample adaption, and offer better computational efficiency for real-time deployment.

When inspecting the results more closely, we find that remaining zero-shot errors are primarily associated with larger-caliber vessels (e.g., in HiP-CT), which are underrepresented in the current generation process. Future work will expand scale diversity within our controllable parameter framework (e.g., increasing initial branch radius) to better model multi-scale vascular hierarchies. Overall, these

findings highlight the feasibility of greatly reducing dependence on large-scale real annotations through structured synthetic data generation while maintaining competitive cross-domain generalization.

5 Conclusion

To address the critical challenge of limited annotated datasets for advancing deep learning-based 3D blood vessel segmentation, we proposed `VesselSim`, a novel two-stage learning framework based on both simulated and real data using unsupervised test-time adaptation. Without needing a single manual segmentation ground truths, the proposed method demonstrated excellent performance across modalities and tasks in comparison with state-of-the-art methods, offering a new direction for data-efficient vascular segmentation.

Acknowledgments. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Fonds de recherche du Québec – Nature et technologies (FRQNT).

Disclosure of Interests. The authors have no competing interests.

References

1. Aktar, M., Rivaz, H., Kersten-Oertel, M., Xiao, Y.: Vesselshot: Few-shot learning for cerebral blood vessel segmentation. In: International Workshop on Machine Learning in Clinical Neuroimaging. pp. 46–55 (2023)
2. Butoi*, V.I., Ortiz*, J.J.G., Ma, T., Sabuncu, M.R., Guttag, J., Dalca, A.V.: Universeg: Universal medical image segmentation. International Conference on Computer Vision (2023)
3. Bériault, S., Xiao, Y., Collins, D.L., Pike, G.B.: Automatic swi venography segmentation using conditional random fields. IEEE Transactions on Medical Imaging **34**(12), 2478–2491 (2015)
4. Colussi, M., Mascetti, S., Dolz, J., Desrosiers, C.: Rec-ttt: Contrastive feature reconstruction for test-time training. In: 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). pp. 6699–6708 (2025)
5. DeVries, T., Taylor, G.W.: Improved regularization of convolutional neural networks with cutout (2017)
6. Falcetta, D., Marciano, V., Yang, K., Cleary, J., Legris, L., Rizzaro, M.D., Pitsiorlas, I., Chaptoukaev, H., Lemasson, B., Menze, B., Zuluaga, M.A.: Vesselverse: A dataset and collaborative framework for vessel annotation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2025. vol. LNCS 15972, pp. 656 – 666. Springer Nature Switzerland (October 2025)
7. Feldman, P., Fainstein, M., Siless, V., Delrieux, C., Iarussi, E.: Vesselvae: Recursive variational autoencoders for 3d blood vessel synthesis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 67–76 (2023)
8. Feldman, P., Sinnona, M., Delrieux, C., Siless, V., Iarussi, E.: Vesselgpt: Autoregressive modeling of vascular geometry. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 662–672 (2025)

9. Galarreta Valverde, M., Macedo, M., Mekkaoui, C.: Three-dimensional synthetic blood vessel generation using stochastic l-systems. In: Medical Imaging 2013: Image Processing. vol. 8669 (03 2013). https://doi.org/10.1117/12.2007532
10. Gu, Y., Sun, Z., Liu, Z., Xu, Y.: Test-time training with local contrast-preserving copy-pasted image for domain generalization in retinal vessel segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 614–624 (2025)
11. Guo, Z., Tan, Z., Feng, J., Zhou, J.: Vesseldiffusion: 3d vascular structure generation based on diffusion model. IEEE Transactions on Medical Imaging (2025)
12. Haast, R.A., Kashyap, S., Ivanov, D., Yousif, M.D., DeKraker, J., Poser, B.A., Khan, A.R.: Insights into hippocampal perfusion using high-resolution, multimodal 7t mri. Proceedings of the National Academy of Sciences **121**(11), e2310044121 (2024)
13. Hamarneh, G., Jassi, P.: Vascusynth: Simulating vascular trees for generating volumetric image data with ground-truth segmentation and tree analysis. Computerized medical imaging and graphics **34**(8), 605–616 (2010)
14. Jain, Y., Borner, K., Walsh, C., Ruschman, N., Lee, P.D., Weber, G.M., Holbrook, R., Howard, A.: Sennet + hoa - hacking the human vasculature in 3d. https://kaggle.com/competitions/blood-vessel-segmentation (2023), kaggle competition
15. Jessen, E., Steinbach, M., Debbaut, C., Schillinger, D.: Branching exponents of synthetic vascular trees under different optimality principles. IEEE transactions on bio-medical engineering **PP** (11 2023). https://doi.org/10.1109/TBME.2023.3334758
16. Moccia, S., De Momi, E., El Hadji, S., Mattos, L.S.: Blood vessel segmentation algorithms — review of methods, datasets and evaluation metrics. Computer Methods and Programs in Biomedicine **158**, 71–91 (2018)
17. Murray, C.D.: The physiological principle of minimum work: I. the vascular system and the cost of blood volume. Proceedings of the National Academy of Sciences **12**(3), 207–214 (1926)
18. Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., Efros, A.: Context encoders: Feature learning by inpainting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2536–2544 (06 2016). https://doi.org/10.1109/CVPR.2016.278
19. PyTorch MONAI Team: Medical open network for ai (MONAI), https://github.com/Project-MONAI/MONAI
20. 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)
21. Sweeney, P.W., Hacker, L., Lefebvre, T.L., Brown, E.L., Gröhl, J., Bohndiek, S.E.: Unsupervised segmentation of 3d microvascular photoacoustic images using deep generative learning. Advanced Science **11**(32), 2402195 (2024)
22. Wang, H., Guo, S., Ye, J., Deng, Z., Cheng, J., Li, T., Chen, J., Su, Y., Huang, Z., Shen, Y., Fu, B., Zhang, S., He, J., Qiao, Y.: Sam-med3d: Towards general-purpose segmentation models for volumetric medical images (2024), https://arxiv.org/abs/2310.15161
23. Wittmann, B., Wattenberg, Y., Amiranashvili, T., Shit, S., Menze, B.: vesselfm: A foundation model for universal 3d blood vessel segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 20874–20884 (2025)

24. Yang, J., Wang, Y.: Design of vascular networks: A mathematical model approach. International Journal for Numerical Methods in Biomedical Engineering **29**(4), 515–529 (2013)
25. Yang, K., Musio, F., Ma, Y., et al.: Benchmarking the cow with the topcow challenge: Topology-aware anatomical segmentation of the circle of willis for cta and mra (2025), https://arxiv.org/abs/2312.17670
26. Yang, Y., Xu, J., Xu, M., Tang, X., Wang, B., Shu, K., You, Z.: Fsvs-net: A few-shot semi-supervised vessel segmentation network for multiple organs based on feature distillation and bidirectional weighted fusion. Information Fusion **123**, 103281 (2025)
27. Yu, W., Jiang, S., Chen, Y., Chang, S., Wang, Y., Wu, B., Dong, J., Liu, M., Zhu, S., Qin, F., Wang, C., Tian, Q.: A large scale benchmark for test time adaptation methods in medical image segmentation. ArXiv **abs/2512.02497** (2025)
28. Zhou, J., Wang, W., Li, S., Qu, X., Guo, X., Liu, Y., Tang, W., Lin, X., Zheng, Y.: Topotta: Topology-enhanced test-time adaptation for tubular structure segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2025)