FreSeg: Frenet-Frame-based Part Segmentation for 3D Curvilinear Structures

Shixuan Gu¹, Jason Ken Adhinarta², Mikhail Bessmeltsev⁴, Jiancheng Yang⁵, Jessica Zhang³, Daniel Berger¹, Jeff W. Lichtman¹, Hanspeter Pfister¹, and Donglai Wei²

Harvard University, USA
 Boston College, USA
 Carnegie Mellon University, USA
 University of Montreal, Canada
 EPFL, Switzerland
 shixuangu@g.harvard.edu, donglai.wei@bc.edu

Abstract. Part segmentation is a crucial task for 3D curvilinear structures like neuron dendrites and blood vessels, enabling the analysis of dendritic spines and aneurysms with scientific and clinical significance. However, their diversely winded morphology poses a generalization challenge to existing deep learning methods, which leads to labor-intensive manual correction. In this work, we propose FreSeg, a framework of part segmentation tasks for 3D curvilinear structures. With a Frenet-Framebased point cloud transformation, it enables the models to learn more generalizable features and have significant performance improvements on tasks involving elongated and curvy geometries. We evaluate FreSeg on 2 datasets: 1) DenSpineEM, an in-house dataset for dendritic spine segmentation, and 2) IntrA, a public 3D dataset for intracranial aneurysm segmentation. Further, we will release the DenSpineEM dataset, which includes roughly 6,000 spines from 69 dendrites from 3 public electron microscopy (EM) datasets, to foster the development of effective dendritic spine instance extraction methods and, consequently, large-scale connectivity analysis to better understand mammalian brains.

Keywords: point cloud reconstruction \cdot 3D segmentation \cdot dendritic spine \cdot medical image dataset \cdot electron microscopy

1 Introduction

Deep learning-enabled 3D biomedical imaging serves both as a catalyst for scientific breakthroughs (e.g., connectomics [5], protein prediction [11]) and as a crucial tool in medical care (e.g., rib analysis [19], aneurysm detection [20]). While the ability of semantic segmentation algorithms to facilitate image analysis and measurement in diverse domains is well-established (e.g., nn-UNet [9]), it remains a challenging task to devise solutions for scenarios with elongated and curvy structures. In clinical research, such generalized cylinders have great

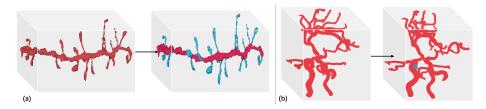


Fig. 1: Part segmentation for 3D curvilinear structures. a) Dendritic spine (blue) segmentation. b) Intracranial aneurysm (blue) segmentation.

biological significance and are common in vasculature systems among different organs and cells (e.g., brains, lungs, retinas). For instance, in neuroscience, the morphology and density of neurites contain critical information about the connectivity strength among neurons [16]. Likewise, in the field of medicine, understanding the shapes and locations of intracranial aneurysms is essential to clinical diagnosis and surgical planning [8]. However, manual inspection and measurement of such structures are labor-intensive for professionals, as generalized cylinder geometries are typically intricate and messy in 3D volumes. Several works have built datasets of such intricate structures to enable the evaluation of deep learning-based methods on such modalities [18,17,20,10].

However, learning robust feature representations on geometries in 3D Euclidean space (e.g., point clouds, volumes, meshes [22,23]) is difficult for deep learning models, since the coordinate systems are sensitive to the symmetries of 3D space, such as rotations and translations [6]. This problem is exacerbated in generalized cylinder geometries which may exhibit symmetries under deformation to its centerlines. Unfortunately, most previous works overlook the geometric complexity inherent in the data representations, and merely devise task-specific methods which lack the capacity to generalize to different domains; for instance, lung vasculature reconstruction tasks model the targets as topological graphs, neglecting the surface geometry which is crucial in other domains such as aneurysm detection. While the self-configuring voxel-based method such as nnU-Net [9] achieved great segmentation results in various biomedical image domains, it failed to embed intricate global geometry or handle dense neurite volumes in high-resolution electron microscopy images (EM images) [18,17]. A few works handle the intricacies of 3D geometry by increasing the expressive capacity of the network with equivariant layers and a high degree of weight sharing [4,6], but the explicit geometry representation is often ignored, which would lead to degradation on extreme scenarios.

In this study, we propose FreSeg, a skeleton-based cylindrical transformation that reconstructs the twisted, elongated, and curvy structure in a more generalizable representation to facilitate deep learning-based tasks. Specifically, FreSeg 1) models the dense 3D volume into sparse point clouds (locally isotropic) to preserve global geometry while reducing computational expense, and 2) reconstructs the point clouds in a cylindrical coordinate system based on the Frenet-Serret Frame (i.e., moving observational frame of reference) along the main skeleton of the target object. With the desirable properties inherent in the transformation

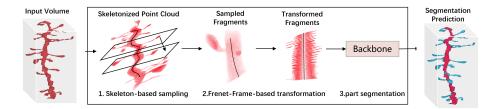


Fig. 2: The Pipeline of FreSeg. The pipeline is divided into 3 steps: 1) Skeleton-based Point Sampling: the dense medical volume is converted to a point cloud for skeletonization [7], and it is further sampled into fragments along the main trunk skeletons. 2) Frenet-Frame-based Transformation: the fragments are reconstructed in a cylindrical coordinate system via a Frenet-Frame-based transformation and propagated into the backbone model for inference. 3) Point Cloud Network Prediction: the final volume prediction is obtained by remapping and combining all predictions with majority voting.

(i.e., bijectivity, differentiability, equivariance), the transformed structures, denoted as *generalized cylinders*, can substantially increase the generalization capacity of the deep learning models. We evaluate the method on two datasets: 1) IntrA, a public 3D dataset for intracranial aneurysm segmentation (reconstructed from MRI images), and 2) **DenSpineEM**, an in-house EM dataset for dendritic spine segmentation, which includes 6,000 spine instances from 69 thoroughly segmented dendrites from three public EM datasets. We conduct comparative experiments of few-shot learning by training the model with and without the transformation under different ratios of the training set to evaluate its capacity for generalization. And by ablation study of influence-time augmentation, we further showed that the rotation-invariant property of the generalized cylinders could generalize the model by decreasing the need for data augmentation. By result analysis, we showed that FreSeg can enable models to learn generalizable features with fewer training samples and has the potential to foster cross-domain learning topics (e.g., transfer learning, synthetic data generation, and few-shot learning) with massive vasculature systems which require huge annotation effort.

Contributions. 1) FreSeg, an equivariant transformation that reconstructs elongated and curvy structures into generalized cylinders, which substantially increases the generalization capacity of deep learning methods. 2) FreSeg has a huge potential to foster cross-domain learning applications in massive vasculature systems. 3) The first large public benchmark for 3D dendritic spine segmentation, which enables downstream applications and method comparison.

2 Methodology

In this study, we proposed *FreSeg*, a framework that 1) samples the curvilinear structure into point cloud fragments based on its skeleton, and 2) utilizes the Frenet-Serret-Frame of the skeleton to reconstruct fragments in a more gen-

eralizable representation to facilitate deep learning-based tasks, as depicted in Fig. 2. To evaluate the method, we applied the framework to tasks in different domains as shown in Fig. 1, *i.e.*, dendritic spine segmentation and intracranial aneurysm segmentation, by modifying the setting of skeleton-based sampling.

2.1 Skeleton-based Point Sampling

Point-based representation. Intuitively, FreSeg reconstructs curvilinear structures in a cylindrical coordinate system to reshape them into tubular structures, as depicted in Fig. 3a. Hence, it models 3D medical imaging as point clouds instead of volume, which also reduces computational expense while preserving the global geometry. Specifically, for DenSpineEM dataset with EM volume, the conversion is simply voxel indexing, while for IntrA dataset with reconstructed artery surfaces, we voxelize the surface [2] and convert it to point clouds.

Skeleton-based sampling. The extreme foreground-background imbalance and data density are common challenges of part segmentation tasks in biomedical domains. To alleviate it, we design a skeleton-based sampling for curvilinear structures which is task-specific. We first skeletonize the point clouds [7] and choose a path as the sampling trajectory. Specifically, for the DenSpineEM dataset where neurites have one trunk connected with multiple spines, we select the path containing most vertices with a degree greater than 2. While for the IntrA dataset where the cerebral artery consists of messy vessels, we traverse the skeleton of the entire artery, prune short branches, and obtain multiple paths to cover the filtered skeleton. And then the point clouds can be easily sampled into fragments, as depicted in Fig. 2.

2.2 Frenet-Frame-based Transformation

Preliminaries. In three-dimensional Euclidean space \mathbb{R}^3 , the Frenet–Serret frame (TNB frame) of a differentiable curve at a point is a triplet of three mutually orthogonal unit vectors (tangent, normal, and binormal). Specifically, let $\mathbf{r}(t)$ be a curve in Euclidean space, representing the position vector of the particle as a function of time, and $\mathbf{s}(t)$ represent the arc length which the particle has moved along the curve in time t [14]. The TNB frame is defined by:

$$\mathbf{T} := \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}, \ \mathbf{N} := \frac{\mathrm{d}\mathbf{T}}{\mathrm{d}s} / \left\| \frac{\mathrm{d}\mathbf{T}}{\mathrm{d}s} \right\|, \ \mathbf{B} := \mathbf{T} \times \mathbf{N}.$$
 (1)

Implementation. Given a point cloud $P = \{(x_i, y_i, z_i) | i = 1, ..., n\}$ and its sampling trajectory $S = \{(x_i, y_i, z_i) | i = 1, ..., m\}$, denoted as the main skeleton, we first calculate the TNB frame of the skeleton according to the definition 2.2, denoted as $TNB_S = \{(t_i, n_i, b_i) | i = 1, ..., m\}$, where t_i , n_i , and b_i are the unit tangent, normal, and binormal vectors of the skeleton S at the point S_i (Fig. 3b). Based on the TNB frame, the point cloud P is reconstructed in a cylindrical coordinate system, denoted as C, where $C = \{(\rho_i, \phi_i, g_i) | i = 1, ..., n\}$.

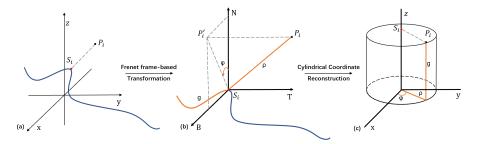


Fig. 3: Calculation of Frenet-Frame-based Transformation. a) For a given point P_i , we find its closest vertex S_i on the curve. b) The orange parts, (ρ, ϕ, g) , represent the calculations based on the Cartesian coordinates and are used as the cylindrical coordinates of the target point P_i , respectively. P'_i is the projection of P_i on the NB plane. c) Point reconstruction in a cylindrical coordinate system.

Specifically, given the point P_i , we first find its closet vertex on the skeleton, denoted as S_i . Assuming S is an ordered set, we have:

$$\rho_i = \operatorname{dis}_{\text{Euclidean}} (P_i, S_j), \quad g_i = \sum_{n=0}^{j-1} \operatorname{dis}_{\text{Euclidean}} (S_n, S_n + 1)$$

$$\phi_i = \begin{cases} \operatorname{arccos} \left(\frac{\boldsymbol{v}_i \cdot \boldsymbol{n}_j}{|\boldsymbol{v}_i| |\boldsymbol{n}_j|} \right), & \text{if } \boldsymbol{v}_i \cdot \boldsymbol{b}_j \ge 0 \\ 2\pi - \operatorname{arccos} \left(\frac{\boldsymbol{v}_i \cdot \boldsymbol{n}_j}{|\boldsymbol{v}_i| |\boldsymbol{n}_j|} \right), & \text{if } \boldsymbol{v}_i \cdot \boldsymbol{b}_j < 0 \end{cases}$$

where v_i presents the projection of the vector $\overrightarrow{S_jP_i}$ (denoted as u_i) on the NB plane, which could be calculated by $v_i = A_jA_j^Tu_i$ ($A_j = [t_j, b_j]$ is a column orthogonal matrix). Fig. 3c gives a visual interpretation of the transformation from Cartesian coordinates to cylindrical coordinates.

Properties. The elegant properties inherent in the transformation (e.g., bijectivity, differentiability, equivariance) enable the model to achieve significant performance improvement and learn more generalizable features. The proof can be found in the supplementary material. (1) **Bijectivity**: with skeleton vertices, the cylindrical representation could be transformed back to the original representation, hence our method does not suffer any information loss. (2) **Rotation invariance**: the representation is consistent regardless of the rotation (and translation) augmentations, hence allowing the model to learn generalizable features.

2.3 Point Cloud Network Prediction

Specifically, for the point cloud fragments with a large scale that is computationally prohibitive, we divide it into equally-sized batches of 30K points. For the batch with insufficient points, we 'ceil' it up by randomly sampling points from other batches, and applying majority voting for repeated points. Finally, the point predictions of all batches are concatenated to obtain the voxel prediction, as depicted in Fig. 2.

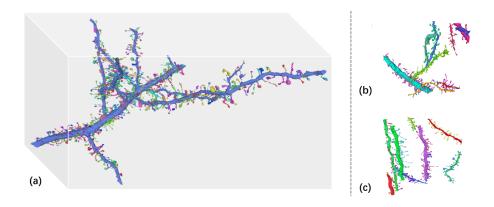


Table 1: **Overview of DenSpineEM Dataset.** We build upon 3 EM volumes with instance segmentation and annotate spine segmentation for 90 dendrites.

Name	Tissue	Size (μm^3)	#Dendrites	#Spines
DenSpine-M50	Mouse Somatosensory Cortex [12]	50x50x30	50	3,886
DenSpine-M	Mouse Visual Cortex [17]	30x30x30	10	315
DenSpine-H	Human Frontal Lobe [17]	30x30x30	9	319

3 DenSpineEM Dataset

The morphology and density of dendritic spines contain critical information about the connectivity strength among neurons. To foster new research to address the spine segmentation challenge, we curated a large-scale 3D dendritic spine segmentation benchmark, DenSpineEM, as depicted in Fig. ??. DenSpineEM contains 6,000 spine instances from around 69 thoroughly segmented dendrites, which consists of 3 3D EM image stacks acquired from mouse (50 μm)³, rat (30 μm)³, and human (30 μm)³ cortical regions, whose neuron segmentation is available from previous publications [12,17]. The development of DenSpineEM took 3 neuroscientists about 30 days in total. Tab. 1 shows the data distribution of the DenSpineEM dataset.

Dataset Construction. We first curate DenSpine-M50 from [12] as our main dataset due its existing segmented dendrites (100+) and spines (4,000+) which are analyzed in [13]. However, the spine segmentation is not thorough on most dendrites, making it hard to train models for practical use due to the unknown false negative errors. Thus, we pick 50 largest dendrites from the existing annotation and manually proofread all spine instance segmentation. In the end, we obtain 3,886 spine instances. We further divide the dendrites into train/val/test split with 30/5/15 dendrites by a rough 60/10/30 ratio for learning-based methods, leading to 2239/512/1135 spine instances respectively. To evaluate the generalization performance of the model trained on DenSpine-M50 across regions

Dataset	DenSpineEM			IntrA* [21]
	M50-test	M	Н	
SO-Net [21]		-/-		71.79
PointNet++ [15]	52.23	29.30	20.10	31.43
FreSeg w/o Trans (Ours)				
FreSeg (Ours)	66.58	42.88	37.84	77.08

Table 2: Quantitative results. The FreSeg is evaluated on two datasets.

and species, we build two small datasets from AxonEM data [17]: DenSpine-M from mouse visual cortex, and DenSpine-H from human frontal lobe.

Annotation Protocol. For a high-quality ground truth annotation, we segment spines manually with the VAST software [3] to avoid introducing bias from automatic methods. To detect errors, we use the Neuroglancer software [1] to generate and visualize 3D meshes of the segmentation of dendrites and spines. Four neuroscience experts were recruited to proofread and double confirm the annotation results for spine instance segmentation.

To evaluate FreSeg, we apply it to both the DenSpineEM and IntrA datasets for part segmentation task. For generalization test, we conducted a comparison study by decreasing training data and applying inference-time augmentation.

3.1 Experiment Setup

DenSpineEM. We trained the model on DenSpine-M50 which contains sufficient samples to be divided into train/val/test, and evaluated the methods on the test set. Since DenSpine-M and DenSpine-H contains much fewer dendrites, we use them as additional test sets to evaluate the generalization performance. For evaluation, we report the dice similarity coefficient (DSC) of the spines.

IntrA. We followed the two-step baseline method [21] to evaluate the methods with five-fold cross-validation, and for a fair comparison with the baseline, we converted point clouds segmentation back to surface models. For generalization evaluation, we compared the performance of models trained with 100%/75%/50%/25% training data scale as well as applying inference-time augmentation (spine-rotation), as shown in Fig. 4. For metrics, we use the DSC of the aneurysm parts for comparison as used in the baseline. The voxelization code, voxelized data, and metrics tables of all 5 folds are in the supplement.

3.2 Benchmark Results

DenSpineEM. We report the metrics for the 3 DenSpineEM datasets with the models trained on the training set of the DenSpineEM-M50 dataset. As reported

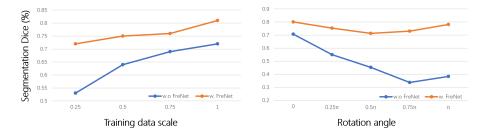


Fig. 4: **Results on IntrA.** To evaluate the generalization capabilities of FreSeg, we report the performance on intracranial aneurysm segmentation vs. a) percentage of training data used and b) inference-time rotation augmentation. The complete 5-fold validation metrics table can be found in the supplement.

in Tab. 2, we showed that FreSeg achieves at 14%/13%/17% higher Dice value than the PointNet++ baseline on the 3 datasets, respectively.

IntrA. As reported in Tab. 2, FreSeg brings a more distinct performance boost on relatively small datasets like IntrA, where it achieves 5% higher Dice value compared to PointNet++, and beats the baseline method by 3%. Qualitative results are provided in the supplementary material.

3.3 Ablation Studies

To further evaluate its capacity of generalization, we conducted an ablation study by decreasing percentage of training data used from 100% to 25%, the method with FreSeg has a much lower performance degradation (-8.5%) than the one without FreNet (-20%) over this range. We also attempt to verify the rotation invariance of FreSeg by applying inference-time rotation augmentations. As reported in Fig. 4b, FreSet maintains relatively stable performance levels (10% drop) while the baseline method suffers a 35% performance drop.

4 Conclusion

In this study, we proposed FreSeg, a framework of part segmentation tasks for 3D curvilinear structures. It reconstructs elongated and curvy geometry to facilitate deep learning applications in (bio)medical scenarios. With the elegant properties inherent in the transformation (e.g., bijectivity, differentiability, equivariance), it enables the model to learn more generalizable features across different domains, which has the potential to foster cross-domain deep learning topics. For instance, FreSeg makes it possible for transfer learning across massive different vasculature datasets to alleviate annotation efforts. In addition, it has the potential to generate synthetic data and facilitate few-shot learning.

References

- 1. Neuroglancer, https://github.com/google/neuroglancer 7
- Barill, G., Dickson, N.G., Schmidt, R.M., Levin, D.I.W., Jacobson, A.: Fast winding numbers for soups and clouds. ACM Transactions on Graphics (TOG) 37, 1 12 (2018) 4
- 3. Berger, D.R., Seung, H.S., Lichtman, J.W.: Vast (volume annotation and segmentation tool): efficient manual and semi-automatic labeling of large 3d image stacks. Frontiers in neural circuits (2018) 7
- 4. Cohen, T., Welling, M.: Group equivariant convolutional networks. In: International Conference on Machine Learning (2016) 2
- Fornito, A., Zalesky, A., Breakspear, M.: The connectomics of brain disorders. Nature Reviews Neuroscience 16, 159–172 (2015)
- Geiger, M., Smidt, T.E.: e3nn: Euclidean neural networks. ArXiv abs/2207.09453 (2022)
- Huang, H., Wu, S., Cohen-Or, D., Gong, M., Zhang, H., Li, G., Chen, B.: L1-medial skeleton of point cloud. ACM Transactions on Graphics (TOG) 32, 1 8 (2013)
 3
- 8. Hunt, W.E., Hess, R.M.: Surgical risk as related to time of intervention in the repair of intracranial aneurysms. Journal of neurosurgery 28 1, 14–20 (1968) 2
- Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods 18, 203–211 (2020) 1, 2
- 10. Jin, L., Gu, S., Wei, D., Kuang, K., Pfister, H., Ni, B., Yang, J., Li, M.: Ribseg v2: A large-scale benchmark for rib labeling and anatomical centerline extraction. ArXiv abs/2210.09309 (2022) 2
- Jumper, J.M., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D.A., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D.: Highly accurate protein structure prediction with alphafold. Nature 596, 583-589 (2021) 1
- Kasthuri, N., Hayworth, K.J., Berger, D.R., Schalek, R.L., Conchello, J.A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T.R., Roberts, M., Morgan, J.L., Tapia, J.C., Seung, H.S., Roncal, W.G., Vogelstein, J.T., Burns, R., Sussman, D.L., Priebe, C.E., Pfister, H., Lichtman, J.W.: Saturated reconstruction of a volume of neocortex. Cell 162(3), 648-661 (Jul 2015). https://doi.org/10.1016/j.cell.2015.06.054
- 13. Ofer, N., Berger, D.R., Kasthuri, N., Lichtman, J.W., Yuste, R.: Ultrastructural analysis of dendritic spine necks reveals a continuum of spine morphologies. Developmental neurobiology (2021) 6
- Piaggio, H.T.H.: Differential geometry of curves and surfaces. Nature 169, 560–560 (1952) 4
- 15. Qi, C., Yi, L., Su, H., Guibas, L.J.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In: NIPS (2017) 7
- 16. Scheffer, L.K., Xu, C.S., Januszewski, M., Lu, Z., Takemura, S.y., Hayworth, K.J., Huang, G.B., Shinomiya, K., Maitlin-Shepard, J., Berg, S., et al.: A connectome and analysis of the adult drosophila central brain. Elife **9**, e57443 (2020) **2**

- 17. Wei, D., Lee, K., Li, H., Lu, R., Bae, J.A., Liu, Z., Zhang, L., Santos, M.d., Lin, Z., Uram, T., et al.: Axonem dataset: 3D axon instance segmentation of brain cortical regions. In: MICCAI (2021) 2, 6, 7
- 18. Wei, D., Lin, Z., Franco-Barranco, D., Wendt, N., Liu, X., Yin, W., Huang, X., Gupta, A., Jang, W.D., Wang, X., et al.: Mitoem dataset: Large-scale 3d mitochondria instance segmentation from em images. In: MICCAI (2020) 2
- 19. Yang, J., Gu, S., Wei, D., Pfister, H., Ni, B.: Ribseg dataset and strong point cloud baselines for rib segmentation from ct scans. In: MICCAI. pp. 611–621. Springer (2021) 1
- 20. Yang, X., Xia, D., Kin, T., Igarashi, T.: Intra: 3d intracranial aneurysm dataset for deep learning. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) pp. 2653–2663 (2020) 1, 2
- Yang, X., Xia, D., Kin, T., Igarashi, T.: A two-step surface-based 3d deep learning pipeline for segmentation of intracranial aneurysms. Computational Visual Media 9, 57–69 (2020) 7
- 22. Zhang, Y.: Challenges and advances in image-based geometric modeling and mesh generation. Image-Based Geometric Modeling and Mesh Generation (2013) 2
- 23. Zhang, Y.J.: Geometric Modeling and Mesh Generation from Scanned Images. CRC Press, Taylor & Francis Group (2016) 2