

Pixel Embedding Method for Tubular Neurite Segmentation

Huayu Fu, Jiamin Li,
Haozhi Qu, Xiaolin Hu, and Zengcai Guo

Tsinghua University
fuhy22@mails.tsinghua.edu.cn

Abstract. Automatic segmentation of neuronal topology is critical for handling large-scale neuroimaging data, as it can greatly accelerate neuron annotation and analysis. However, the intricate morphology of neuronal branches and the occlusions among fibers pose significant challenges for deep learning-based segmentation. To address these issues, we propose an improved framework:

First, we introduce a deep network that outputs pixel-level embedding vectors and design a corresponding loss function, enabling the learned features to effectively distinguish different neuronal connections within occluded regions. Second, building on this model, we develop an end-to-end pipeline that directly maps raw neuronal images to SWC-formatted neuron structure trees. Finally, recognizing that existing evaluation metrics fail to fully capture segmentation accuracy, we propose a novel topological assessment metric to more appropriately quantify the quality of neuron segmentation and reconstruction.

Experiments on our fMOST imaging dataset demonstrate that, compared to several classical methods, our approach significantly reduces the error rate in neuronal topology reconstruction. The code and dataset have been released in https://github.com/FullIsCool/embed_net.

Keywords: fiber occlusion segmentation · neuron structure tree reconstruction · pixel embedding · connectivity evaluation

1 Introduction

Automatic segmentation and three-dimensional reconstruction of individual neurons are key to deciphering the brain’s microscale connectome and functional networks. The morphology of a single neuron’s dendrites and axon, together with the spatial distribution of its synapses, determines the neuron’s input–output patterns. High-precision morphological reconstructions enable the assembly of detailed brain connectomes, thereby revealing the topological organization of sensory, motor, and cognitive circuits [14,12].

Fluorescence Micro-Optical Sectioning Tomography (fMOST) [25] combines precision mechanical sectioning with high-sensitivity fluorescence imaging to achieve whole-brain, single-cell resolution ($0.35 \times 0.35 \times 1.0 \mu\text{m}$ /voxel). On platforms such as Wuhan OE-Bio BioMap5000, fMOST is highly automated, yet

a single mouse brain can generate on the order of 16TB of image data—far exceeding manual tracing capabilities. Even with semi-automated tools [11], annotation throughput lags behind data acquisition, making labeling the primary bottleneck for large-scale quantitative analysis.

Recently, deep learning-based methods for automated neuron reconstruction have emerged [30,1,22]. However, fully automatic, high-fidelity reconstructions remain elusive due to two fundamental challenges. Figure 1 illustrates these two challenges First, fiber crossing and self-entanglement: dendrites and axons often wind around themselves or intersect with processes from other neurons over long distances, a scenario distinct from standard 3D instance-segmentation occlusions. Overlapping branches in the image volume frequently lead two-class semantic segmentation networks to merge unrelated fibers, producing fused boundaries and topological errors. Second, boundary ambiguity and low signal-to-noise ratio: under fMOST, neurite diameters span only a few pixels and are affected by optical diffraction, uneven fluorescence labeling, and background noise, yielding gray-scale transition zones several pixels wide. Such blurred edges impede accurate segmentation and often induce breakpoints or spurious connections during skeleton extraction.

These limitations restrict the applicability of existing approaches to whole-brain datasets and prevent reconstructions from meeting the dual demands of accuracy and completeness in neuroanatomical and connectomic studies.

To address these challenges, we propose an end-to-end reconstruction framework based on pixel-level embedding. Our network outputs a high-dimensional feature vector for each voxel, so that distances in embedding space naturally separate different fiber segments and avoid the branch-merging pitfalls of binary segmentation. A subsequent graph-based post-processing module further refines connections at crossing and discontinuity sites. Recognizing the critical importance of connectivity in neuronal morphology, we also introduce a multi-scale connectivity evaluation metric that, unlike a single Betti-number measure, quantifies connection errors at various scales. Experiments using this metric demonstrate that our method substantially reduces connectivity errors.

Our main contributions are as follows:

1. We reformulate neuron reconstruction as a pixel embedding problem and design a discriminative embedding loss tailored to fiber separation.
2. We develop an end-to-end reconstruction pipeline that converts raw fMOST volumes directly into SWC-format neuron trees.
3. We propose a graph-topology-driven connectivity metric that quantifies connection errors at multiple scales, enhancing the biological credibility of reconstructed neurons.

2 Related Work

3D Neuron Reconstruction Methods. Since the release of V3D (now Vaa3D) by Peng et al. in 2010[18], this platform has become a leading tool for large-

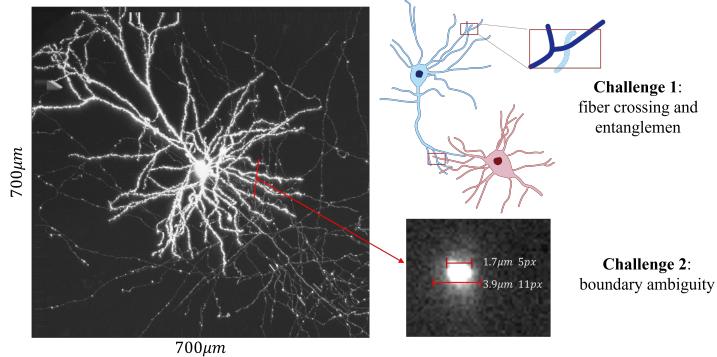


Fig. 1: challenges

scale 3D bioimage visualization and analysis. Vaa3D supports smooth, real-time browsing of terabyte-scale, multi-resolution datasets, fully meeting the demands of whole-brain, high-resolution visualization. Building on Vaa3D, the V3D-Neuron plugin has introduced various automatic tracing algorithms. For example, the APP2 (All-Path Pruning) algorithm employs a gray-weighted distance transform combined with front propagation to assemble tree-like topologies. The UltraTracer method, grounded in maximum-likelihood estimation, completes reconstruction through statistical modeling[19]. In 2019, Li et al.[15] proposed the G-Cut algorithm, which leverages quantified neurite growth orientation features and a graph-search strategy to provide a globally optimal segmentation of overlapping neuronal branches.

Manual, expert-guided annotation remains the gold standard for accurate neuron morphology reconstruction. Engineering optimizations have significantly improved annotation efficiency. In 2017, Gou et al.[10] developed the Fast Neuron Tracer (FNT), which integrates the entire tracing process into a cohesive software suite aligned with human annotation logic, resulting in the largest corpus of 6,357 mouse prefrontal cortex projection neurons to date[9]. More recently, the Gapr software[11] introduced further enhancements, including multi-user, multi-device annotation support and U-Net-based AI assistance.

Deep Learning-Based Methods. Ronneberger et al.[21] first introduced the U-Net architecture to address biomedical segmentation tasks characterized by small foregrounds and stringent accuracy requirements. Çiçek et al.[4] extended U-Net to volumetric data (3D U-Net). Subsequent works have proposed variants such as U-Net++[29], Attention U-Net[16], and ResUNet[7], enhancing feature extraction through revised backbones, refined skip connections, or attention modules.

Given the tubular geometry of vessels and neurites, specialized convolutional operations have been explored: Yu et al.[28] employed dilated convolutions to enlarge the receptive field; Dai et al.[6] introduced deformable convolutions to

adaptively capture geometric deformations; Qi et al.[20] designed dynamic snake convolutions for smooth curve modeling. These modules were integrated into U-shaped backbones to create networks such as D-UNet[13], DeUNet[8], and DCU-net[27], demonstrating marked improvements on public datasets of retinal and myocardial vasculature.

Other studies have focused on loss functions tailored for tubular structure segmentation. Shit et al.[23] proposed the cLDICE similarity measure to strengthen tubular recognition. Wang et al.[26] introduced Deep Distance Transform (DDT), a geometry-aware segmentation framework. While these methods emphasize continuity constraints, inaccuracies and offsets in the extracted skeleton can undermine their effectiveness. Leveraging persistent homology, King et al.[5] formulated a topological loss function that enforces topological similarity during training to restore continuity in fragmented regions.

3 Methodology

3.1 Pixel embedding U-Net

Based on the 3D U-Net architecture of Ronneberger et al. [4], we modify the final layer to produce an n -dimensional embedding for each voxel. Specifically, we replace the original single-channel $1 \times 1 \times 1$ convolution with a $1 \times 1 \times 1$ convolution having n output channels. The remainder of the encoder-decoder backbone retains its four-level symmetric U-shape, thereby preserving 3D U-Net’s strength in foreground/background binary segmentation.

Our model adopts a dual-branch output:

Segmentation branch: maintains the original foreground/background classification task, trained with a binary cross-entropy loss to enforce accurate encoding and decoding.

Embedding branch: outputs an n -dimensional feature vector for each voxel and employs a metric-learning loss to encourage tight clustering of vectors within the same fiber segment and clear separation between different segments.

As illustrated in Fig.2(a), the network implements a differentiable mapping

$$\begin{aligned} f_\theta: \mathbb{R}^{H \times W \times D} &\longrightarrow \mathbb{R}^{n \times H \times W \times D}, \\ \mathbf{X} = f_\theta(\mathbf{V}), \quad \mathbf{X} \in \mathbb{R}^{n \times H \times W \times D}, \quad \mathbf{x}_i \in \mathbb{R}^n &\quad \text{for each voxel } i. \end{aligned} \tag{1}$$

where \mathbf{v}_i denotes the i -th input voxel and \mathbf{x}_i is its corresponding n -dimensional embedding. The complete set of outputs $\{\mathbf{x}_i\}$ thus provides highly discriminative representations for subsequent neuron segmentation and connectivity analysis.

3.2 Discriminative Loss Based on Local Vector Clustering

To effectively separate interleaved and overlapping neurite segments in the high-dimensional embedding space, we adapt the discriminative loss of De Brabandere et al.[3] and incorporate the long-range continuity and local occlusion characteristics of neuronal fibers. The key idea is: first, Use prior ground-truth labels to

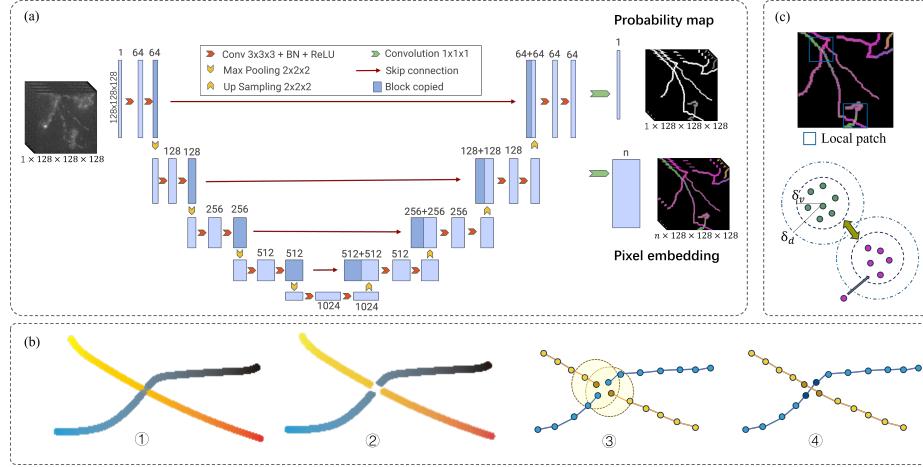


Fig. 2: (a)Schematic diagram of the model structure using pixel embedding. (b)post-processing for generating SWC. (c)Embedding vector clustering within a local field of vision

locate small patches where overlaps occur. Then, Within each patch, treat each micro-segment as an independent instance and impose clustering constraints on its embedding vectors.

Our local discriminative loss comprises four terms:

- (i) Variance term L_{var} (intra-cluster compactness)
- (ii) Distance term L_{dist} (inter-cluster separation)
- (iii) Continuity term L_{con} (smoothness along fibers)
- (iv) Regularization term L_{reg} (embedding norm penalty)

Let δ_v and δ_d be the intra-cluster and inter-cluster margins, respectively. We denote by C_k the number of instances in local patch k , by S_c the set of voxels belonging to instance c , and by μ_c the centroid of those voxels' embedding vectors. The variance term measures the distance between each voxel's embedding and its instance centroid, enforcing $\|x_i - \mu_c\| \leq \delta_v$ to ensure adequate intra-cluster compactness. The distance term ensures that centroids of different instances are well separated in the embedding space by requiring $\|\mu_c - \mu_{c'}\| \geq 2\delta_d$. As is explained in Fig2(c).

$$L_{\text{var}}^{(k)} = \frac{1}{C_k} \sum_{c=1}^{C_k} \frac{1}{|S_c|} \sum_{i \in S_c} [\|x_i - \mu_c\| - \delta_v]_+^2, \quad (2)$$

$$L_{\text{dist}}^{(k)} = \frac{1}{C_k(C_k - 1)} \sum_{\substack{c_A, c_B=1 \\ c_A \neq c_B}}^{C_k} [2\delta_d - \|\mu_{c_A} - \mu_{c_B}\|]_+^2, \quad (3)$$

where $\|\cdot\|$ denotes the Euclidean norm and $[x]_+ = \max(0, x)$. The superscript k indexes patches.

To enforce smooth variation along each neurite, let $A(x_i)$ be the set of embedding vectors in the 26-neighborhood of x_i . The continuity term is

$$L_{\text{con}} = \frac{1}{2|\{x_i\}|} \sum_{x_i} \frac{1}{|A(x_i)|} \sum_{x'_i \in A(x_i)} [\|x'_i - x_i\| - \delta_v]_+^2. \quad (4)$$

To prevent feature explosion, we add a regularization term:

$$L_{\text{reg}} = \frac{1}{|\{x_i\}|} \sum_{x_i} \|x_i\|^2. \quad (5)$$

During training, we also include the binary cross-entropy loss L_{bce} from the segmentation head. The total loss is

$$L_{\text{total}} = \frac{1}{N_{\text{patch}}} \sum_{k=1}^{N_{\text{patch}}} (\alpha L_{\text{var}}^{(k)} + \beta L_{\text{dist}}^{(k)}) + \gamma L_{\text{con}} + \eta L_{\text{reg}} + \xi L_{\text{bce}}, \quad (6)$$

where $\alpha, \beta, \gamma, \eta, \xi$ are weighting hyperparameters and N_{patch} is the total number of local patches. This design ensures that crossing fibers are locally well separated while maintaining global smoothness and compactness of each neurite's embedding.

3.3 Post-processing methods for establishing the neural fiber skeleton

In neuro-morphological studies, the skeleton representation encapsulates the essential topological information of neuronal processes. Specifically, a series of nodes are placed along each neurite, and their parent-child relationships define a tree structure. The SWC (Standardized Wireframe Container) format [2] is widely adopted for quantitative description of 3D neuron morphology. Each SWC record comprises seven fields: node index, type, three-dimensional coordinates, radius, and parent node index, thereby fully encoding the spatial connectivity of the neuron. Compared to detailed geometric meshes, the skeleton more directly reflects topological features such as branching patterns and hierarchy, making it especially valuable for morphometric analyses.

To reconstruct neurons from the learned embeddings, we first apply our re-trained segmentation network to obtain a binary foreground mask and compute an n -dimensional embedding vector for every voxel. Next, we impose an embedding-distance threshold ϵ on each pair of adjacent foreground voxels: if the Euclidean distance between their embeddings exceeds ϵ , the link is deemed a cross-fiber break. This criterion partitions the foreground mask into a collection of micro-segments. We then perform 3D parallel thinning [17] on these segments to extract a one-voxel-wide skeleton.

To restore continuity across true neurite crossings and small gaps, we examine each skeleton node’s neighborhood: any two adjacent skeleton voxels whose embedding distance falls below ϵ are reconnected. Finally, a breadth-first search (BFS) on the resulting connectivity graph yields a complete tree structure. Figure 2(c) illustrates the entire pipeline from embedding field to final skeleton, including the “jump-connection” strategy at crossing points.

4 Experiments

4.1 Connectivity Evaluation Algorithm

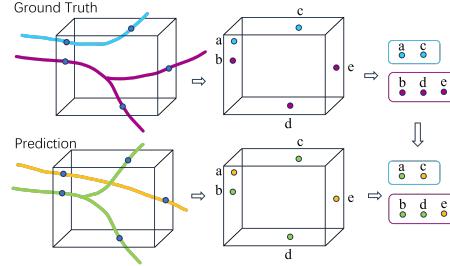


Fig. 3: Schematic diagram of connectivity evaluation algorithm

In neuronal connectomics, our focus lies not on microscopic branch details but on the projection directions and global connectivity of neural fibers. As illustrated in Figure 3, two reconstructions may share the same Betti number yet exhibit markedly different connection patterns—differences that a single scalar topological metric cannot capture.

To address this, we propose a multiscale connectivity evaluation pipeline:

1. **End-Point Extraction.** On both the ground-truth and predicted SWC trees, identify all nodes located at volume boundaries or branch termini. These constitute the terminal node sets.

2. **Ground-Truth Component Partition.** Using the connectivity graph of the ground-truth SWC tree, partition its terminal nodes into connected subsets. Each subset represents the projection of a single neurite at the boundary. Similarly, partition the predicted terminals according to the predicted SWC connectivity.

3. **Terminal Pairing.** Establish a one-to-one correspondence between predicted and ground-truth terminals via nearest-neighbor matching in Euclidean space.

4. **Error Classification.** – *Type I (Disconnection).* If terminals from a single ground-truth component are paired with terminals in two or more distinct predicted components, this indicates a false break—i.e., a true connection has been

erroneously severed. – *Type II (Overconnection)*. If a single predicted component contains terminals paired to two or more different ground-truth components, this indicates a spurious merge—i.e., fibers that should remain separate are incorrectly connected.

By tallying these two error types, we achieve a fine-grained, scale-aware quantification of connectivity fidelity, overcoming the limitations of traditional single-value topological metrics.

4.2 Experimental Results

Datasets.

We constructed an in-house fMOST neuronal imaging dataset. Ai166 reporter-line mice were crossed with a layer-specific Cre driver to sparsely label pyramidal neurons in layers I–V of the prefrontal cortex [24]. Whole-brain imaging was then performed on the BioMap5000 fMOST system (Wuhan OE-Bio) at $0.35 \times 0.35 \times 1.0 \mu\text{m}$ voxel resolution [25]. From three mice, we randomly extracted 180 subvolumes exhibiting strong fluorescence, capturing somata, dendrites, and axons. Using the Fast Neuron Tracer (FNT) software [10], expert annotators manually placed skeleton nodes within each subvolume. Each region was then partitioned into 128^3 -voxel blocks and augmented by random rotations and mirror flips, yielding a dataset of 33,000 fully annotated 128^3 samples for training and evaluation.

Baseline Models.

To assess the topological fidelity afforded by our embedding network, we compared against three standard U-Net variants: the vanilla 3D U-Net, U-Net++ [29], and Attention U-Net [16]. All networks shared an identical channel configuration to ensure fair comparison. This study tests the hypothesis that improvements in pure binary segmentation accuracy alone cannot deliver comparable gains in topological reconstruction without embedding-based representations.

Implementation Details.

We monitor performance along three axes: (1) Segmentation quality—Accuracy, Precision, Recall, and F1 score; (2) Overlap metrics—Dice coefficient and Intersection over Union (IoU); (3) Topological connectivity—our proposed skeleton-connectivity evaluation. The dataset was split 80%/20% into training and validation sets. All models were trained with the Adam optimizer, an initial learning rate of 1×10^{-4} , and identical remaining hyperparameters.

Results.

Table 1 summarizes segmentation and overlap metrics for the four models. Since our embedding network builds upon the same 3D U-Net backbone for foreground extraction, its segmentation metrics match those of the vanilla 3D U-Net. U-Net++ underperforms on our dataset, while Attention U-Net achieves improvements of 5.2% in Dice and 3.4% in IoU. However, because the ground-truth

masks are derived from densely annotated SWC skeletons, even the baseline U-Net already attains near-ideal segmentation—hence further gains offer limited benefits for downstream skeleton extraction.

Table 3 reports the connectivity error counts. For Type I (disconnection) errors, all well-segmented models register counts near zero, whereas U-Net++ hampered by segmentation breaks exhibits a higher Type I rate. For Type II (overconnection) errors, our embedding network cuts the error count by nearly 50% compared to the best baseline. Overall, without increasing computational complexity or parameter count, our method substantially reduces both disconnection and overconnection errors.

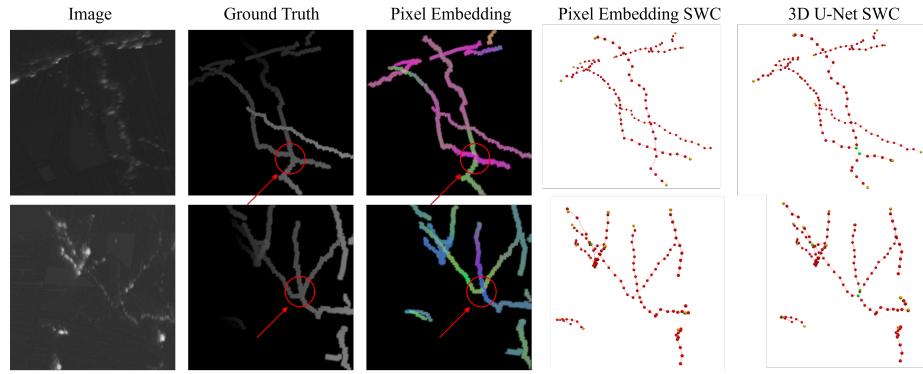


Fig. 4: Results demonstration. First column: raw input image (the actual input is a 3D volume; shown here as a 2D projection). Second column: ground truth — to visualize the 3D structure, different depths are rendered in varying shades of gray. Third column: predicted vector field, with high-dimensional embeddings mapped into color space. Fourth column: SWC node connectivity predicted by our method. Fifth column: SWC connectivity predicted by the U-Net baseline; mutually connected branch points are highlighted in green.

Table 1: Comparison of Models’s Semantic Segmentation Performance

	Accuracy	Precision	Recall	F1-score	Dice	IoU
embed	99.93%	85.74%	95.35%	85.53%	0.8553	0.8170
3D U-Net	99.93%	85.74%	95.35%	85.53%	0.8553	0.8170
U-Net++	99.94%	82.89%	74.25%	73.76%	0.7376	0.6208
Att U-Net	99.98%	91.22%	92.84%	90.76%	0.9076	0.8515

Table 2: Comparison of Models’ Connection Performance

	correct	typeI	typeII	typeI/correct	typeII/correct
embed	15.2	0.03	1.62	2.0×10^{-3}	0.106
3D U-Net	13.7	0.006	3.18	4.4×10^{-4}	0.232
U-Net++	13.6	3.3	1.97	0.24	0.145
Att U-Net	14.3	0.006	2.56	4.2×1^{-4}	0.179

Table 3: Model parameter quantity and inference time (running on NVIDIA A800 80 GB Tensor Core GPU)

	Params	Inference time(s)
embed	90.4M	0.054
3D U-Net	90.3M	0.039
U-Net++	104.8M	0.152
Att U-Net	91.4M	0.068

5 Conclusion

This study systematically optimizes connectivity errors in automatic neuron reconstruction under fMOST imaging. We first introduce a deep network that outputs pixel-level embedding vectors and design a discriminative loss to ensure that, even in regions of overlap and occlusion, different neuronal fibers are effectively separated in the high-dimensional feature space. Building on this model, we develop an end-to-end reconstruction pipeline that converts raw volumetric data directly into SWC-formatted skeleton trees. To overcome the limitations of traditional metrics in assessing structural fidelity, we propose a novel topological evaluation metric that quantifies segmentation and reconstruction quality at multiple scales and dimensions. Experimental results demonstrate that our method significantly reduces connectivity errors. Nevertheless, in exceedingly complex fiber-entanglement scenarios, current U-Net-based architectures remain insufficient, underscoring the need for more innovative algorithms to further enhance reconstruction robustness.

6 Acknowledgment

References

1. Callara AL, Magliaro C, Ahluwalia A, and Vanello N. A smart region-growing algorithm for single-neuron segmentation from confocal and 2-photon datasets. *Front Neuroinform*, 14, Mar 2020.
2. Giorgio A. Ascoli, David E. Donohue, and Maryam Halavi. NeuroMorpho.Org: A central resource for neuronal morphologies. *Journal of Neuroscience*, 27(35):9247–9251, 2007.
3. Bert De Brabandere, Davy Neven, and Luc Van Gool. Semantic instance segmentation with a discriminative loss function. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 7–9, 2017.
4. Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: Learning dense volumetric segmentation from sparse annotation. In Sébastien Ourselin, Leo Joskowicz, Mert R. Sabuncu, Gozde Unal, and William Wells, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, pages 424–432, Cham, 2016. Springer International Publishing.
5. James R. Clough, Nicholas Byrne, Ilkay Oksuz, Veronika A. Zimmer, Julia A. Schnabel, and Andrew P. King. A topological loss function for deep-learning based image segmentation using persistent homology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(12):8766–8778, 2022.
6. Jifeng Dai, Haozhi Qi, Yuwen Xiong, Yi Li, Guodong Zhang, Han Hu, and Yichen Wei. Deformable convolutional networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 764–773, 2017.
7. Foivos I. Diakogiannis, François Waldner, Peter Caccetta, and Chen Wu. Resunet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162:94–114, 2020.
8. Shunjie Dong, Zixuan Pan, Yu Fu, Qianqian Yang, Yuanxue Gao, Tianbai Yu, Yiyu Shi, and Cheng Zhuo. Deu-net 2.0: Enhanced deformable u-net for 3d cardiac cine mri segmentation. *Medical Image Analysis*, 78:102389, 2022.
9. Le Gao, Sang Liu, Lingfeng Gou, Yachuang Hu, Yanhe Liu, Li Deng, Danyi Ma, Haifang Wang, Qiaoqiao Yang, Zhaoqin Chen, Dechen Liu, Shou Qiu, Xiaofei Wang, Danying Wang, Xinran Wang, Biyu Ren, Qingxu Liu, Tianzhi Chen, Xiaoxue Shi, Haishan Yao, Chun Xu, Chengyu T. Li, Yangang Sun, Anan Li, Qingming Luo, Hui Gong, Ninglong Xu, and Jun Yan. Single-neuron projectome of mouse prefrontal cortex. *Nature Neuroscience*, 25(4):515–529, Apr 2022.
10. Liangjie Gou. Fast neurite tracer (version 1.0). GNU Free Documentation License v1.2+, 2017. Version 1.0, released 2017-02-24.
11. Lingfeng Gou, Yanzhi Wang, Le Gao, Yiting Zhong, Lucheng Xie, Haifang Wang, Xi Zha, Yinqi Shao, Huatai Xu, Xiaohong Xu, and Jun Yan. Gapr for large-scale collaborative single-neuron reconstruction. *Nature Methods*, 21(10):1926–1935, 2024.
12. Seung HS. *Connectome: how the brain's wiring makes us who we are*. Houghton Mifflin Harcourt, Boston, 2012.
13. Qiangguo Jin, Zhaopeng Meng, Tuan D. Pham, Qi Chen, Leyi Wei, and Ran Su. Dunet: A deformable network for retinal vessel segmentation. *Knowledge-Based Systems*, 178:149–162, 2019.
14. Lichtman JW and Sanes JR. Ome sweet ome: what can the genome tell us about the connectome? *Current Opinion in Neurobiology*, 18(3):346–353, 2008.

15. Rui Li, Muye Zhu, Junning Li, Michael S. Bienkowski, Nicholas N. Foster, Hanpeng Xu, Tyler Ard, Ian Bowman, Changle Zhou, Matthew B. Veldman, X. William Yang, Houri Hintiryan, Junsong Zhang, and Hong-Wei Dong. Precise segmentation of densely interweaving neuron clusters using g-cut. *Nature Communications*, 10(1):1549, 2019.
16. Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018.
17. Kàlmàn Palàgyi and Attila Kuba. A 3d 6-subiteration thinning algorithm for extracting medial lines. *Pattern Recognition Letters*, 19(7):613–627, 1998.
18. Hanchuan Peng, Zongcai Ruan, Fuhui Long, Julie H. Simpson, and Eugene W. Myers. V3D enables real-time 3d visualization and quantitative analysis of large-scale biological image data sets. *Nature Biotechnology*, 28(4):348–353, Apr 2010.
19. Hanchuan Peng, Zhi Zhou, Erik Meijering, Ting Zhao, Giorgio A Ascoli, and Michael Hawrylycz. Automatic tracing of ultra-volumes of neuronal images. *Nature Methods*, 14(4):332–333, 2017.
20. Yaolei Qi, Yuting He, Xiaoming Qi, Yuan Zhang, and Guanyu Yang. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 6047–6056, 2023.
21. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.
22. Li S, Quan T, Zhou H, Huang Q, Guan T, Chen Y, Xu C, Kang H, Li A, Fu L, Luo Q, Gong H, and Zeng S. Brain-wide shape reconstruction of a traced neuron using the convex image segmentation method. *Neuroinformatics*, 18(2):199–218, 2020.
23. Suprosanna Shit, Johannes C. Paetzold, Anjany Sekuboyina, Ivan Ezhov, Alexander Unger, Andrey Zhylka, Josien P. W. Pluim, Ulrich Bauer, and Bjoern H. Menze. cldice - a novel topology-preserving loss function for tubular structure segmentation. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 16555–16564, 2021.
24. Marijke B. Veldman, Christine S. Park, Charles M. Eyermann, et al. Brainwide genetic sparse cell labeling to illuminate the morphology of neurons and glia with cre-dependent MORF mice. *Neuron*, 108:111–127.e116, 2020.
25. Xiaojun Wang, Hanqing Xiong, Yurong Liu, Tao Yang, Anan Li, Fei Huang, Fangfang Yin, Lei Su, Ling Liu, Ning Li, Longhui Li, Shenghua Cheng, Xiaoxiang Liu, Xiaohua Lv, Xiuli Liu, Jun Chu, Tonghui Xu, Fuqiang Xu, Hui Gong, Qingming Luo, Jing Yuan, and Shaoqun Zeng. Chemical sectioning fluorescence tomography: high-throughput, high-contrast, multicolor, whole-brain imaging at subcellular resolution. *Cell Reports*, 34(5):108709, 2021.
26. Yan Wang, Xu Wei, Fengze Liu, Jieneng Chen, Yuyin Zhou, Wei Shen, Elliot K. Fishman, and Alan L. Yuille. Deep distance transform for tubular structure segmentation in ct scans. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3832–3841, 2020.
27. Xin Yang, Zhiqiang Li, Yingqing Guo, and Dake Zhou. Dcu-net: a deformable convolutional neural network based on cascade u-net for retinal vessel segmentation. *Multimedia Tools and Applications*, 81(11):15593–15607, May 2022.

28. Fisher Yu, Vladlen Koltun, and Thomas Funkhouser. Dilated residual networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 636–644, 2017.
29. Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. In Danail Stoyanov, Zeike Taylor, Gustavo Carneiro, Tanveer Syeda-Mahmood, Anne Martel, Lena Maier-Hein, João Manuel R.S. Tavares, Andrew Bradley, João Paulo Papa, Vasileios Belagiannis, Jacinto C. Nascimento, Zhi Lu, Sailesh Conjeti, Mehdi Moradi, Hayit Greenspan, and Anant Madabhushi, editors, *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 3–11, Cham, 2018. Springer International Publishing.
30. zhou Z, Kuo HC, Peng H, and Long F. Deepneuron: an open deep learning toolbox for neuron tracing. *Brain Informatics*, 5:3, 2018.