Improved Topological Preservation in 3D Axon Segmentation and Centerline Detection using Geometric Assessment-driven Topological Smoothing (GATS)

Nina I. Shamsi Alec S. Xu Lars A. Gjesteby Laura J. Brattain

MIT Lincoln Laboratory, Lexington, MA 02421, USA

n.shamsi@ece.neu.edu, alecx@umich.edu

{lars.gjesteby, brattainl}@ll.mit.edu*

Abstract

Automated axon tracing via fully supervised learning requires large amounts of 3D brain imagery, which is time consuming and laborious to obtain. It also requires expertise. Thus, there is a need for more efficient segmentation and centerline detection techniques to use in conjunction with automated annotation tools. Topology-preserving methods ensure that segmented components maintain geometric connectivity, which is especially meaningful for applications where volumetric data is used, and these methods often make use of morphological thinning algorithms as the thinned outputs can be useful for both segmentation and centerline detection of curvilinear structures. Current morphological thinning approaches used in conjunction with topology-preserving methods are prone to over-thinning and require manual configuration of hyperparameters.

We propose an automated approach for morphological smoothing using geometric assessment of the radius of tubular structures in brain microscopy volumes, and apply average pooling to prevent over-thinning. We use this approach to formulate a loss function, which we call Geometric Assessment-driven Topological Smoothing loss, or GATS. Our approach increased segmentation and centerline detection evaluation metrics by 2%-5% across multiple datasets, and improved the Betti error rates by 9%. Our ablation study showed that geometric assessment of tubular structures achieved higher segmentation and centerline detection scores, and using average pooling for morphological smoothing in place of thinning algorithms reduced the Betti errors. We observed increased topological preservation during automated annotation of 3D axons volumes from models trained with GATS.

1. Introduction

Curvilinear structures are line-like objects with differences in pixel intensities relative to neighboring pixels [1]; a line is a 1-dimensional manifold, though a curvilinear structure may not necessarily be 1-dimensional [11]. Curvilinear structure segmentation is segmenting binary masks of curvilinear structures [1]. Methods that ensure maximizing geometric connectivity of segmented curvilinear structures are said to be topology-preserving [2, 5, 25], and are used for performing topologically accurate segmentation. A variety of domains and applications require topologically accurate segmentation: 3D axon tracing and centerline detection [19], retinal vessel segmentation [27], and airway tree reconstruction [4, 34], to name a few.

Skeletonization-based approaches for curvilinear structure segmentation are useful for both segmentation and centerline detection [2, 25]. Morphological thinning algorithms for volumetric data have previously utilized neighborhood lookup tables [20] or directional sub-iteration [33], but such algorithms can be computationally expensive [17, 29]. The centerline Dice (clDice) loss function is an example of a topology-preserving method that has shown to be performant for curvilinear structure segmentation and centerline detection, especially with respect to model training time, through a process called soft-skeletonization [25]. However, parameterizing soft-skeletonization is currently done via educated guesswork [19, 23, 25].

This paper presents a novel approach to determine the number of iterations to perform thinning or smoothing operations using the mean pixel radius of axon segments from N random slices of volumetric data. We use this method to automatically parameterize morphological thinning or smoothing algorithms. Moreover, we observed that certain affine rotation transformations would degrade the performance of models which utilized skeletonization, and so we modified our approach to use morphological smoothing via average

pooling operations instead of using morphological thinning via min-, max- pooling. We show that our approach, Geometric Assessment-driven Topological Smoothing (GATS), increased segmentation and centerline detection metrics by $2\% - 5\%$ across multiple datasets with 3D axon imagery. Moreover, topology-preserving methods can be used for automated annotation of unannotated volumes of brain imagery [10, 19]. The present work shows an application of GATS for automated annotation of 3D brain imagery. As far as we know, our work is the first to present a topology-preserving morphological smoothing approach for curvilinear structure segmentation and centerline detection in 3D brain imagery. Therefore, the contributions of this work are as follow:

1. a morphological smoothing method for curvilinear structures using the mean pixel radius of tubular structures across N random slices from a volumetric input and average pooling operations.

2. automated annotation of axon segments in 3D brain imagery using our approach, which prevents over-thinning of axons and thus promotes topological preservation in automatically annotated volumes.

2. Related Literature

2.1. Topology-Preserving Losses for Curvilinear Structure Segmentation

Presently the literature for curvilinear structure segmentation and centerline detection focuses on thinning algorithms, so we compare our approach to thinning and non-thinning methods. However, morphological smoothing has been utilized for 3D images [17], and biomedical images previously [14, 28, 32]. Our research aims to maximize topology-preservation of segmented axons from 3D volumes as axons can present with various lengths and a single axon can span across an entire input volume. Moreover, as we are interested in automated annotation of unannotated brain imagery, we seek a method that can precisely and accurately detect an axon's centerline.

2.1.1 Centerline Dice (clDice)

clDice was developed to measure the topology-preservation of tubular or curvilinear structure segmentation because other measures of segmentation quality, e.g. Dice, would report the same segmentation quality for models with comparably different performance for segmenting an input volume with both small/fine and large/coarse tubular structures [25], i.e., the degree of a model's segmentation quality was not reflected in the scoring metric. The clDice metric reflects the degree of topology preserved following segmentation by comparing the intersection of masks and skeletons using

Algorithm 1 Soft Skeletonization

procedure SKEL(I, k) ▷ k is set manually
  $I' \leftarrow maxpool(minpool(I))$
  $S \leftarrow ReLU(I - I')$
  for `i to k` **do**
    $I \leftarrow minpool(I)$
    $I' \leftarrow maxpool(minpool(I))$
    $Delta \leftarrow ReLU(I - I')$
    $S \leftarrow S + ReLU(Delta - S \circ Delta)$
  end for
  return S
end procedure

measures of topological precision (Eq. 1) and topological sensitivity (Eq. 2), where the harmonic mean of Eq. 1 and Eq. 2 results in a clDice score (Eq. 3); a higher score indicates a higher degree of structural connectivity. S_P is the predicted skeleton, and V_L is the ground truth volume; similarly, S_L is the ground truth skeleton, and V_P is the predicted segmentation mask. The process of computing clDice is an algorithm called soft-clDice.

$$Tprec(S_P, V_L) = \frac{|S_P \cap V_L|}{|S_P|} \tag{1}$$

$$Tsens(S_L, V_P) = \frac{|S_L \cap V_P|}{|S_L|} \tag{2}$$

Skeletonization is the process of iteratively removing foreground pixels in a binary image till a construct remains that maintains the extent and connectivity of the original object [3]. The soft skeletonization algorithm (Algorithm 1), which produces the skeletons used to compute topological precision and topological sensitivity, performs morphological thinning of a curvilinear structure by applying iterative min- and max- pooling using a data-dependent hyperparameter *k*. The clDice paper (2021) notes that the *k* hyperparameter must be greater than the pixel radius of the largest observed tubular structure in a given dataset.

$$clDice(V_P, V_L) = 2 \times \frac{Tprec(S_P, V_L) \times Tsens(S_L, V_P)}{Tprec(S_P, V_L) + Tsens(S_L, V_P)} \tag{3}$$

Soft-skeletonization, via soft-clDice, produces a differentiable result that can be used for neural network training, which is important because usually morphological thinning is a discrete operation and produces a non-differentiable result [25, 34]. clDice provides a balance for preserving topology while learning segmentation of curvilinear structures by setting $\alpha \in [0, 0.5]$; an α greater than 0.5 may favor learning skeleta [25]. The clDice method relies on extracting accurate skeletons, and as such the performance of the method depends on hyperparameter tuning for an optimal

k value for a given dataset, which if done manually can be onerous and likely to find a suboptimal solution [24]. Additionally, the iterative process of min- and max- pooling used during soft skeletonization may result in loss of features that are critical to the topology of the structure.

2.1.2 Homotopy Warping

The homotopic warping algorithm functions on the principle that given two binary masks with the same topology, one mask can be warped into the other mask by sequential flipping of simple points [5]. The connectivity of a pixel *p* on a 2D binary image, or a voxel *v* on a 3D binary volume, is defined by its neighboring pixels/voxels. The homotopy warping algorithm identifies topological critical points on a given binary mask (e.g., segmentation prediction) by warping it into another given binary mask (e.g., segmentation ground truth), and a resultant mask with the topological critical pixels, i.e., non-simple points, is used to compute a homotopy warping error-based loss where the critical pixels denote topological errors in warping one mask to the other. A distance transform is proposed by Hu [5] for optimal pixel flipping, and thus identification of critical pixels. Compared to clDice, the homotopic warping algorithm does not appear to increase Dice or adjusted Rand Index (ARI) scores as much as would be desired given the increase in training time [5]. However, the distance-ordered homotopy warping method can be used outside of the context of the homotopic warping loss function, and as such is useful as a generalizable method for identification of topologically critical points.

3. Proposed Method

3.1. Morphological Smoothing using Mean Pixel Radius and Average Pooling

As mentioned our research aim is to minimize a topology-preserving loss to precisely and accurately segment axons from 3D brain imagery, and find the centerlines of their curvilinear structures for automated annotation. As such we decided to formulate a loss function for this study. Our proposed method automatically determines the number of iterations required for a smoothing algorithm using geometric assessment of input data. The goal is to prevent over-thinning and loss of fine/small curvilinear structure features.

3.1.1 Mean Pixel Radius for N Random Slices

A tubular structure can be represented by its boundaries and the cross-sectional radius of each point in its skeleton [4, 9, 31], and so the maximum radius is the greatest distance from the medial axis of a skeleton point. For soft-skeletonization, using a *k* parameter smaller than the largest pixel radius of a tubular structure in a dataset may result in incomplete thinning [25]. We propose that for morphological approaches that rely on knowing the pixel radii of tubular structures in a given volumetric input, we may necessarily only need to know the maximum radius in a volumetric input with multiple tubular structures of varying cross-sectional radii. The maximum pixel radii of tubular structures from *N* random slices of a given volumetric input can be averaged, and the result from that computation can be used as the number of iterations required to perform erosion and dilation operations for a thinning or smoothing algorithm.

Algorithm 2 Mean Pixel Radius

procedure MPR(I) ▷ I is a 3D, 4D, or 5D input
  $d \leftarrow 0$
  $s \leftarrow Slice(I)$
  for `slice in s` **do**
    $c \leftarrow Canny(slice)$
    $Dist \leftarrow MedialAxis(c)$
    $MaxDist \leftarrow max(Dist)$
    $d \leftarrow d + (2 \times MaxDist)$
  end for
  return $int(2 \times (\frac{d}{len(s)}))$ ▷ of N random slices
end procedure

Algorithm 3 Topological Smoothing

procedure TS(I, k) ▷ k is predetermined using **MPR**
  $I' \leftarrow avepool(avepool(I))$
  $S \leftarrow ReLU(I - I')$
  for `i to k` **do**
    $I \leftarrow avepool(I)$
    $I' \leftarrow avepool(avepool(I))$
    $Delta \leftarrow ReLU(I - I')$
    $S \leftarrow S + ReLU(Delta - S \circ Delta)$
  end for
  return S
end procedure

Given a binary image *I*, let *d* be the mean pixel radius, initialized as zero. The *Slice* function reshapes the 3D input and chooses without replacement, *N* random slices, forming the resultant list of 2D inputs, *s*. The boundary of tubular structures is determined using a Canny edge detector to enable computing the medial axis of the structure, and the distance to the edge boundary from the medial axis is returned by the *MedialAxis* function from scikit-image. We experimentally found that returning twice the maximum distance and twice the mean pixel radius was more robust with our random sampling approach. Thus, the MPR method, Algorithm 2, returns a geometrically determined integer for estimating the iterations of thinning or smoothing for a given volumetric input with varying cross-sectional widths of tubular structures.

3.1.2 Topological Smoothing

Some skeletonization algorithms produce constructs that are sensitive to rotations due to directional bias [26, 31, 33], so we decided not to use a skeletonization approach for topology-preservation. Some of our initial experiments show that soft-skeletonization appears sensitive to data-dependant rotation transformations, likely due to increased morphological thinning at some rotations relative to others. We include these initial experiments in the Supplementary Material, but they form the basis for why we chose to use morphological smoothing. Following the logic of soft-skeletonization, we instead apply average pooling operations to the intermediate constructs of the algorithm, where we aim to induce topological smoothing by iteratively subtracting a more open construct from a less open construct; see line 7 in seen in Algorithm 3. Using this method, we aim to produce an output that is less prone to over-thinning. For Algorithm 3, k is the number of iterations to perform the smoothing algorithm. We evaluate our approach on its ability to maintain geometrical connectivity as measured by metrics indicating topological preservation.

3.2. Loss Function Formulations

Similarly to clDice, we use the harmonic mean of the overlap of the masks and smoothed outputs,

$$GATS(V_P, V_L) = 2 \times \frac{Tprec(T_P, V_L) \times Tsens(T_L, V_P)}{Tprec(T_P, V_L) + Tsens(T_L, V_P)} \tag{4}$$

where T_P is the predicted smoothed output, and V_L is the ground truth volume; similarly, T_L is the ground truth smoothed output, and V_P is the predicted segmentation mask. We believe that the overlap of smoothed output and ground truth still gives a relative measure of topological precision and topological sensitivity. We used the objective function in Eq. 5 for training models with GATS, where α was fixed at 0.5. We formulated variants of the GATS loss function to systematically determine the effect of each component: 1) the MPR algorithm, 2) the smoothing algorithm, and 3) GATS. We reasoned that if we use the MPR algorithm with min-, max- pooling we can still achieve soft skeletonization of the inputs, and so we compared using the MPR method with both min-, max- pooling which enables soft skeletonization (GASK), and average pooling (GATS). We did an ablation study using our different loss functions.

$$L_G = (1 - \alpha)(1 - \mathbf{Dice}) + \alpha(1 - \mathbf{GATS}) \tag{5}$$

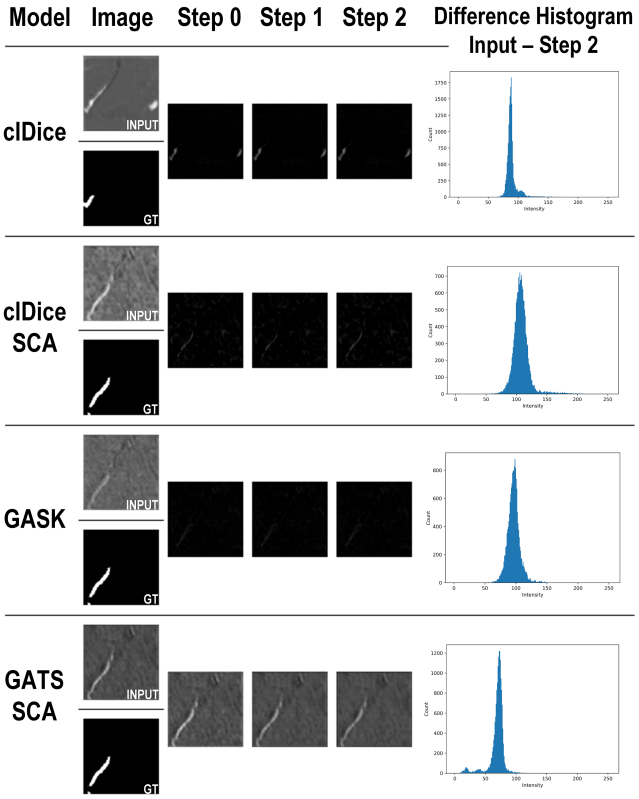

Figure 1. **Morphological thinning progress on DS2.** The best performing models on DS2 are compared. Iterations of soft-skeletonization (GASK) or topological smoothing (GATS), which are automatically determined by MPR while manually set for clDice, are both three. Therefore, three steps of skeletonization/smoothing are shown. The difference histogram between the input and Step 2 shows difference in pixel intensities, where the smoothed output difference results in intensities centered around lower ranges, indicating that the input and Step 2 output are closer in pixel intensities. This is expected for a smoothing operation. SCA: spatial and channel attention. GT: ground truth.

4. Experimental Methodology

4.1. Datasets

As we are interested in automated annotation of unannotated volumes of axons, we conducted our experiments using 3D brain imagery data.

4.1.1 Dataset 1

Dataset 1 (DS1) is a light sheet microscopy dataset of 3X expanded mouse brain tissue, where a stain was used to target Parvalbumin positive neurons from the globus pallidus externus (PVGPe). The CLARITY method was used to stabilize the tissue with clear hydrogels, preserving biomolecules and enabling removal of lipids, which makes the unstained portions of sample optically transparent [12].

Table 1. **Mean and standard deviation evaluated on the test set of Dataset 1 (DS1) and Dataset 2 (DS2) across 10 trials. For clDice: $\alpha = 0.65, k = 3$, GASK and GATS both use MPR, for which N=10, GASK uses soft-skeletonization, GATS uses topological smoothing (TS). For TS: k=5 for DS1, k=3 for DS2.**

Dataset	Attention	Model	Dice ↑	clDice ↑	ρ-Dice ↑	Adj. Rand ↑	β_0 Error ↓	β_1 Error ↓
DS1	None	*clDice*	0.793 ± 0.007	0.744 ± 0.022	0.737 ± 0.017	0.584 ± 0.014	0.348 ± 0.018	0.066 ± 0.008
		GASK	0.790 ± 0.008	0.710 ± 0.029	0.725 ± 0.038	0.580 ± 0.017	0.291 ± 0.033	0.035 ± 0.011
		GATS	0.800 ± 0.007	0.757 ± 0.016	0.761 ± 0.013	0.598 ± 0.014	0.316 ± 0.019	0.061 ± 0.008
	SCA	*clDice*	0.801 ± 0.011	0.761 ± 0.026	0.759 ± 0.025	0.600 ± 0.022	0.334 ± 0.016	0.074 ± 0.014
		GASK	0.790 ± 0.012	0.701 ± 0.027	0.718 ± 0.031	0.579 ± 0.024	**0.270 ± 0.039**	0.074 ± 0.013
		GATS	**0.819 ± 0.006**	**0.799 ± 0.012**	**0.798 ± 0.009**	**0.636 ± 0.011**	0.314 ± 0.022	**0.060 ± 0.007**
DS2	None	*clDice*	0.788 ± 0.014	0.689 ± 0.027	0.775 ± 0.029	0.576 ± 0.027	0.127 ± 0.031	0.151 ± 0.017
		GASK	0.698 ± 0.018	0.741 ± 0.013	0.804 ± 0.011	0.395 ± 0.036	0.024 ± 0.009	0.370 ± 0.017
		GATS	**0.803 ± 0.007**	0.715 ± 0.016	**0.806 ± 0.016**	**0.605 ± 0.014**	0.047 ± 0.010	**0.096 ± 0.014**
	SCA	*clDice*	0.725 ± 0.008	**0.746 ± 0.008**	0.799 ± 0.007	0.450 ± 0.016	0.091 ± 0.023	0.367 ± 0.016
		GASK	0.691 ± 0.019	0.738 ± 0.025	0.796 ± 0.023	0.381 ± 0.038	**0.021 ± 0.004**	0.374 ± 0.008
		GATS	0.787 ± 0.011	0.682 ± 0.021	0.772 ± 0.024	0.573 ± 0.021	0.050 ± 0.012	0.104 ± 0.021

Table 2. **Slicing experiments for MPR (Algorithm 2). Mean and Standard Deviation of Evaluation Metrics for Differently Sliced Variants of GATS on the Test Set of Dataset 1 and Dataset 2 across 10 trials, For clDice: $\alpha = 0.65, k = 3$. For TS: k=5 for DS1, k=3 for DS2. Time indicates average minutes for a single trial (n=10).**

Dataset	Model	Dice ↑	clDice ↑	ρ-Dice ↑	Adj. Rand ↑	β_0 Error ↓	β_1 Error ↓	Time ↓
DS1	clDice	0.793 ± 0.007	0.744 ± 0.022	0.737 ± 0.017	0.584 ± 0.014	0.348 ± 0.018	0.066 ± 0.008	**26.627 ± 7.676**
	GATS N=2	0.803 ± 0.006	0.756 ± 0.014	0.762 ± 0.015	0.605 ± 0.012	0.339 ± 0.010	0.066 ± 0.003	38.507 ± 11.088
	GATS N=3	0.803 ± 0.007	0.761 ± 0.014	0.767 ± 0.015	0.605 ± 0.013	**0.321 ± 0.021**	**0.058 ± 0.007**	46.307 ± 13.314
	GATS N=4	**0.808 ± 0.004**	**0.773 ± 0.009**	**0.778 ± 0.010**	**0.615 ± 0.008**	0.334 ± 0.024	0.061 ± 0.006	38.398 ± 11.050
	GATS N=5	0.803 ± 0.006	0.765 ± 0.014	0.768 ± 0.017	0.604 ± 0.012	0.331 ± 0.017	0.066 ± 0.009	42.945 ± 12.348
DS2	clDice	0.788 ± 0.014	0.689 ± 0.027	0.775 ± 0.030	0.576 ± 0.027	0.127 ± 0.031	0.151 ± 0.017	**30.723 ± 8.845**
	GATS N=2	0.795 ± 0.007	0.700 ± 0.016	0.789 ± 0.015	0.589 ± 0.016	**0.045 ± 0.010**	**0.094 ± 0.012**	34.483 ± 9.939
	GATS N=3	0.796 ± 0.006	0.701 ± 0.015	0.793 ± 0.013	0.591 ± 0.011	0.052 ± 0.009	0.102 ± 0.012	35.299 ± 10.168
	GATS N=4	0.800 ± 0.008	**0.710 ± 0.021**	**0.801 ± 0.020**	0.599 ± 0.016	0.049 ± 0.009	0.101 ± 0.015	36.480 ± 10.512
	GATS N=5	**0.800 ± 0.006**	0.709 ± 0.014	0.800 ± 0.013	**0.599 ± 0.013**	0.051 ± 0.005	0.106 ± 0.009	34.787 ± 10.018

Table 3. **Ablation study mean and standard deviation evaluated on the test set of Dataset 1 (DS1) and Dataset 2 (DS2) across 10 trials. MPR is Mean Pixel Radius, Smooth is Topological Smoothing, Soft Skel is soft-skeletonization**

Ablated	Dataset	Model	Dice ↑	clDice ↑	ρ-Dice ↑	Adj. Rand ↑	β_0 Error ↓	β_1 Error ↓
MPR	DS1	*Soft Skel (k=5)*	0.794 ± 0.010	0.707 ± 0.023	0.718 ± 0.031	0.586 ± 0.020	0.253 ± 0.058	0.070 ± 0.015
		Smooth (k=5)	0.810 ± 0.009	0.782 ± 0.017	0.781 ± 0.015	0.618 ± 0.018	0.326 ± 0.036	0.065 ± 0.010
	DS2	*Soft Skel (k=5)*	0.749 ± 0.008	0.715 ± 0.019	0.801 ± 0.006	0.497 ± 0.017	0.095 ± 0.038	0.303 ± 0.020
		Smooth (k=5)	0.801 ± 0.010	0.710 ± 0.019	0.794 ± 0.018	0.601 ± 0.019	0.053 ± 0.010	0.102 ± 0.019
Smooth	DS1	*No Attn*	0.804 ± 0.007	0.756 ± 0.012	0.768 ± 0.015	0.607 ± 0.014	0.262 ± 0.066	0.057 ± 0.016
		SCA	0.798 ± 0.032	0.734 ± 0.089	0.740 ± 0.090	0.594 ± 0.064	0.187 ± 0.074	0.046 ± 0.016
	DS2	*No Attn*	0.812 ± 0.001	0.735 ± 0.001	0.826 ± 0.001	0.623 ± 0.001	0.083 ± 0.005	0.151 ± 0.004
		SCA	0.792 ± 0.060	0.695 ± 0.012	0.785 ± 0.014	0.584 ± 0.012	0.095 ± 0.009	0.147 ± 0.008
GATS	DS1	*Dice*	0.805 ± 0.011	0.763 ± 0.038	0.769 ± 0.038	0.609 ± 0.023	0.196 ± 0.066	0.048 ± 0.020
	DS2	*Dice*	0.776 ± 0.002	0.663 ± 0.003	0.751 ± 0.004	0.551 ± 0.003	0.076 ± 0.011	0.119 ± 0.008

The PVGPe volume in its entirety is 2048×2048×1271 voxels, and has a voxel resolution of 0.6×0.6×2 µm, where only a 256×256×206 voxel (148×148×412 µm) subvolume was manually annotated [19]. Model training was done on 128×128×64 sized voxel samples using a contiguous subdivided volumes of DS1 in a 50:25:25 training, test and validation split.

Table 4. **Betti error mean and standard deviation evaluated on the test set of Janelia across 10 trials. All models use SCA attention. For clDice: $\alpha = 0.65, k = 3$. For Smooth Dice and GATS: N=4.**

Architecture	Model	β_0 Error ↓	β_1 Error ↓
3DResSE UNet	*clDice*	0.614 ± 0.159	0.047 ± 0.005
	Warping	0.716 ± 0.003	0.049 ± 0.001
	Dice	0.685 ± 0.005	0.052 ± 0.002
	Smooth Dice	0.709 ± 0.178	0.048 ± 0.013
	GATS	**0.200 ± 0.002**	**0.036 ± 0.003**
Cascading 3D UNet	*clDice*	0.641 ± 0.004	0.050 ± 0.001
	Warping	0.600 ± 0.001	**0.049 ± 0.001**
	Dice	0.700 ± 0.012	0.050 ± 0.006
	Smooth Dice	**0.394 ± 0.177**	0.052 ± 0.001
	GATS	0.631 ± 0.001	0.051 ± 0.001

4.1.2 Dataset 2

Dataset 2 (DS2) consists of 20X magnified samples from the mouse thalamus which were labeled via cortical injection with recombinant adeno-associated virus expressing tdTomato (red) and synaptophysin (green). Imagery was then acquired using a Leica confocal microscope. The tdTomato channel was converted to grayscale for use in this paper. The cross-section of this data volume is 581.250 μm^2, thickness is 35 µm, lateral pixel resolution is 0.142 μm^2, and axial resolution is 0.69 µm. The full DS2 volume is 4096×4096×52 voxels. The training, testing, and validation volumes were split in the same ratios as those for DS1, except the voxel size samples were 128×128×32 for model training.

4.1.3 Janelia

We also conducted experiments using the Janelia dataset from the BigNeuron Project [18], which consists of optical microscopy data of single neurons from the adult Drosophila nervous system. Janelia has 42 volumes of data, and so we allocated 30 volumes for training, six for validation and six for testing. We ended up using only one volume for testing in case more training data was needed. All volumes were scaled between zero and one using min-max normalization and we used crop size of 128×128×32 for model training.

4.2. Model Implementation and Training

Our experiments on DS1 and DS2 consisted of training a Residual 3D UNet with four resolution blocks using one of our formulated loss functions and comparing performance against clDice [25]. Our experiments on Janelia consisted of using a Residual 3D UNet with squeeze and excitation blocks (3DResSE UNet), and a Cascading 3D UNet as previously described [19] with four and three resolution blocks for the voxel-wise segmentation head and centerline detection head, respectively. The Cascading 3D UNet used a multi-input loss formulated as:

$$L_M = (1-\alpha)(1-\textbf{segloss}) + \alpha(1-\textbf{clloss}) \quad (6)$$

where α=0.8, and *seg loss* (segmentation loss) was the chosen loss function for the voxel-wise segmentation task, and *cl loss* (centerline loss) was the chosen loss function for the centerline prediction task. We also compared performance on the Janelia dataset against Warping loss [5]. Warping loss supersedes both TopoNet loss [6] and DMT loss [7], and so we only tested against Warping loss. We tried GATS with N=2, 3, 4, 5, or 10 while profiling the average training time required for the 10 trials. Optimal hyperparameters for clDice were determined experimentally by testing α=0.5 or 0.65, and *k*=3 or 8. All our models were trained using an Intel Xeon G6 node (40 cores) with 2 NVIDIA Volta V100 GPUs of 337 GiB memory.

We used a PyTorch framework for data processing, algorithm development and model training. All data used were pre-processed by clipping the highest and lowest 0.01% of values, applying a median filter, and scaling between 0 and 1. Besides using data augmentation during training and inference as previously described in [19], we used a series of affine rotation transformations that were particular for each dataset. Each experiment was repeated 10 times using a model with or without 3D spatial and channel attention (SCA) [8]. We tried efficient channel attention [30] and triplet attention [16] as well (Supplementary Material), but found that evaluation metrics were lower than those for models trained on SCA, so only SCA results are reported.

The 3D U-Net implementation consisted of a 3×3×3 convolution layer followed by group normalization and activation using exponential linear units. Strided 2×2×2 max-pooling and strided transpose convolutions with max pooling were used for downsampling and upsampling, respectively. An ADAM optimizer coupled with cosine annealing was used, with an initial learning rate of 1×10^{-4} and weight decay of 1×10^{-3}. Each trial consisted of randomly cropped and augmented samples from the input data in mini batches of 16 samples. The segmentation output from each model was skeletonized using a 3D-skeletonization algorithm to acquire single voxel wide centerlines. Models were evaluated using the Dice coefficient for voxel-wise similarity [35], the clDice metric [25] and Betti number errors for topology preservation [22], ρ-Dice coefficient for centerline detection accuracy [19], and the adjusted Rand Index (ARI) for ground truth and predicted clusterings equivalence [21]. The Betti numbers β_0 and β_1 measure the number of distinct connected components and circular holes, respectively [25]. All tables bold the best metric in each category.

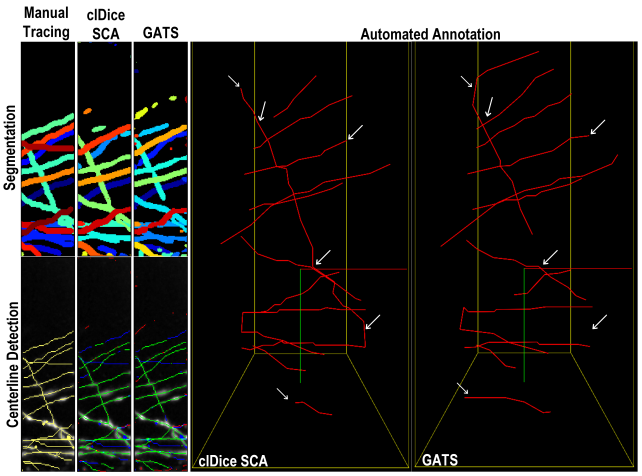

Figure 2. **Comparison of 3D Axon Projections for DS1.** The segmentation results (top) are contrasted with the centerline detection results (bottom), with automated annotation of a 3D volume (right). For centerline detection results, true positives (green), false negatives (blue), and false positives (red) are shown. The red label for false positive does not apply to the automated annotation results. The white arrows show differences in axonal connectivity across models.

5. Results

5.1. Axon Segmentation and Centerline Detection

We compared GATS against two topology-preserving losses, one which uses skeletonization (clDice) and one which does not (Warping). The performance of Warping loss was only compared for the Janelia dataset as even after 24 hours Warping loss does not finish training three of 10 trials on DS1 or DS2, which does not make a reasonable comparison for either clDice or our method.

For results in Table 1, MPR algorithmically determined its own number of iterations to perform either soft skeletonization (GASK) or topological smoothing (GATS). A comparison of thinning or smoothing outputs is shown (Figure 1). MPR (Algorithm 2) determined k=5 for DS1 and k=3 for DS2. For DS1 we found that when the number of iterations used to perform skeletonization are acquired via the MPR method, the model performance drops unless topological smoothing is performed (GATS), as seen in Table 1 and Figure 2, indicating the need for morphological smoothing for topological preservation. The GASK models perform poorer on DS1 because MPR determined k=5 leads to over-thinning with a method like soft-skeletonization, but not with the topological smoothing method. For DS2 GASK without attention is comparable in performance to clDice with manually selected hyperparameters, indicating the efficacy of using geometric assessment for determining the optimal number of iterations for morphological thinning (see Figure 3 for a visual comparison). On DS2 the GATS model performs better for all metrics except the centerline detection metric (clDice) than the model trained on the clDice loss as alpha is affixed to 0.5 for GASK/GATS models. This shows the sensitivity of alpha and k parameter choices for segmentation and centerline detection. Ideally, hyperparameter tuning can be data driven and not manually optimized, as GASK/GATS aims to do. The MPR algorithm can select any N random slices (without replacement) from a given input volume, and the training time of N = 2, 3, 4, 5 and 10 was compared. We found that when N=10, the mean training time for 10 trials with GATS was double that of the mean training time of trials with clDice (5.2h vs. 10h). However, choosing N=4 seemed sufficient for each dataset, though slower by 1h for DS2 and 2h for DS1 (Table 2); this is most likely due to the use of the Canny detector for boundary determination. Based on our evaluation, it would appear that adding an attention mechanism is useful for datasets with lower voxel resolution like DS1.

5.2. Ablation Study

We carried out an ablation study by systematically removing one of three components from GATS in three separate trials using models without attention unless indicated otherwise. The first trial consisted of removing *MPR* and manually setting k=5 as it was previously determined by MPR as the optimal number of iterations for topological smoothing for DS1 but not DS2. The second trial consists of removing the average pooling-based morphological thinning and using a general thinning algorithm from Scikit Image to perform morphological thinning while using MPR to determine the number of iterations to perform. The third trial consisted of removing GATS and using a Dice loss. For DS1, GASK without attention is comparable to MPR-ablated Soft Skel (k=5), where k=5 was the number of iterations previously determined by MPR as optimal for DS1. GATS without attention has a worse evaluation metrics than MPR-ablated Smooth(k=5), though the latter has worse Betti error rates, which is unexpected but maybe likely due to the random sampling approach for MPR which gave an incorrect assessment for at least one of the ten trials used to compute the average for GATS without attention. However, given that MPR-ablated k=5 model with topological smoothing (Smooth, k=5) on DS1 performs better than clDice without attention (0.782 vs. 0.744), there is merit to finding the number of iterations required for a thinning or smoothing algorithm as trials in Table 1 use k=3 for DS1. On DS2 the MPR-ablated model (k=5) with topological smoothing performs marginally worse than its GATS counterpart without attention (Table 1, row 9 vs. Table 3, row 4). The same holds for the MPR-ablated soft-skeletonization model without attention, though its Dice, ARI and Betti 1 scores are marginally better for the MPR ablated version. The attention model for DS1 without topological smoothing but

with a general thinning algorithm for skeletonization has a worse centerline line detection metric (clDice) versus GATS with attention (0.735 vs. 0.799), indicating that for DS1 the average-pooling based topological smoothing was effective for preserving topological connectivity. However the smoothing-ablated model has marginally less Betti errors relative to both clDice and GATS, indicating the need to explore different approaches for morphological thinning or smoothing. For DS2, removing attention, topological smoothing and using a general thinning algorithm results in a more performant model (Table 1, row 9 vs. Table 3, row 7) relative to the dice metrics used for comparison. This indicates that for inputs with higher voxel resolution topological smoothing might be deleterious, while these inputs may still benefit from topological skeletonization (Table 1, row 8 vs. Table 3, row 7).

5.3. Comparing Topology-Preservation on Janelia

For automated annotation of unannotated brain imagery, a simpler solution may be to improve curvilinear structure segmentation using a topology preserving loss without any multitasking for centerline detection. We compared the performance of topology-preserving losses via their Betti errors on two architectures: 1) 3DResSE UNet, and 2) a multi-headed Cascading 3D UNet as described earlier [19]. MPR determined k=3 for Janelia. In Table 4, Smooth Dice was given inputs that were passed through topological smoothing. Based on the best scores for the models trained on the loss functions we compared, we find that topological smoothing reduces Betti error rates for the Janelia dataset.

5.4. Automated Annotation of Unannotated Brain Imagery

To visualize automatic annotation of unannotated 3D brain imagery, we used NeuroTrALE [15], a variant of Neuroglancer, which is a WebGL-based viewer for volumetric data. We used the best performing clDice or GASK/GATS trained model according to the clDice metric on DS1 and DS2. As mentioned, Warping loss does not finish training even three of 10 trials on DS1 or DS2 in 24 hours, and so was not used as a comparison on these datasets. Models trained on our method produce longer and less disconnected axons, as seen in Figure 2 for DS1 and Figure 3 for DS2. For the centerline detection results in Figures 2 and 3, green indicates true positive, blue indicates false negative and red indicates false positive for model predictions relative to the ground truth. Manual tracings are the ground truth.

6. Discussion

Topology-preserving methods which use skeletonization may over-thin 3D brain imagery, and so techniques such as average pooling-based smoothing, as seen in our topological smoothing method, may be needed. However, our ablation study suggests that the approach used by both soft-skeletonization and topological smoothing is possibly too aggressive for the type of 3D brain imagery used here, even though a loss function formulated using topological smoothing performs relatively better for segmentation and centerline detection. We find it limiting that our best models are all some flavor of a UNet architecture, and aim to incorporate our findings into other deep learning architectures, e.g., like an adapter module for Segment Anything [13].

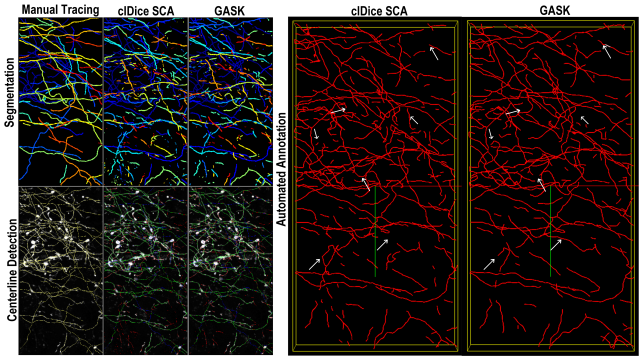

Figure 3. **Comparison of 3D Axon Projections for DS2.** 3D axon predictions of DS2 show that the GASK version produces longer and more connected axon segments (white arrows), which shows that soft-skeletonization can overthin if an incorrect k hyperparameter is chosen. For centerline detection results, true positives (green), false negatives (blue), and false positives (red) are shown. The red label for false positive does not apply to the automated annotation results.

7. Conclusive Remarks

We show that the optimal number of thinning or smoothing iterations for a morphological operation can be determined using N random slices for a volumetric input of brain imagery, and that average pooling-based morphological smoothing can improve both segmentation and centerline detection metrics, indicating increased topological preservation. We also show that our loss function, GATS, can be used to train models for automatic annotation of volumes of brain imagery.

8. Acknowledgement

The authors acknowledge Kwanghun Chung at MIT for providing data, and MIT Lincoln Laboratory Supercomputing Center for their support of high performance computing tasks. This work was supported by the National Institutes of Health (NIH) U01MH117072, NIMH 1R43MH128076-01, and MBF Bioscience. Any opinions, conclusions, or recommendations in this material are those of the author(s) and do not necessarily reflect the views of NIH/NIMH.

References

[1] P. Bibiloni, M. González-Hidalgo, and S. Massanet. A survey on curvilinear object segmentation in multiple applications. *Pattern Recognition*, 60:949–970, 2016. 1

[2] Mingfei Cheng, Kaili Zhao, Xuhong Guo, Yajing Xu, and Jun Guo. Joint topology-preserving and feature-refinement network for curvilinear structure segmentation. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 7127–7136, 2021. 1

[3] Ching-Shoei Chiang. *The Euclidean distance transform*. Purdue University, 1992. 2

[4] Florent Grélard, Fabien Baldacci, Anne Vialard, and Jean-Philippe Domenger. New methods for the geometrical analysis of tubular organs. *Medical Image Analysis*, 42:89–101, 2017. 1, 3

[5] Xiaoling Hu. Structure-aware image segmentation with homotopy warping, 2022. 1, 3, 6

[6] Xiaoling Hu, Fuxin Li, Dimitris Samaras, and Chao Chen. Topology-preserving deep image segmentation. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. 6

[7] Xiaoling Hu, Yusu Wang, Li Fuxin, Dimitris Samaras, and Chao Chen. Topology-aware segmentation using discrete morse theory. In *International Conference on Learning Representations*, 2021. 6

[8] Mobarakol Islam, V. S. Vibashan, V. Jeya Maria Jose, Navodini Wijethilake, Uppal Utkarsh, and Hongliang Ren. Brain tumor segmentation and survival prediction using 3d attention unet. In Alessandro Crimi and Spyridon Bakas, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 262–272, Cham, 2020. Springer International Publishing. 6

[9] Bertrand Kerautret, Adrien Krähenbühl, Isabelle Debled-Rennesson, and Jacques-Olivier Lachaud. 3D Geometric Analysis of Tubular Objects based on Surface Normal Accumulation. In *18th International Conference on Image Analysis and Processing*, Genova, Italy, Sept. 2015. 3

[10] Tzofi Klinghoffer, Peter Morales, Young-Gyun Park, Nicholas Evans, Kwanghun Chung, and Laura J. Brattain. Self-supervised feature extraction for 3d axon segmentation, 2020. 2

[11] T.M. Koller, G. Gerig, G. Szekely, and D. Dettwiler. Multiscale detection of curvilinear structures in 2-d and 3-d image data. In *Proceedings of IEEE International Conference on Computer Vision*, pages 864–869, 1995. 1

[12] Taeyun Ku, Justin Swaney, Jeong-Yoon Park, Alexandre Albanese, Evan Murray, Jae Hun Cho, Young-Gyun Park, Vamsi Mangena, Jiapei Chen, and Kwanghun Chung. Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues. *Nature biotechnology*, 34(9):973–981, 2016. 4

[13] Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, , and Bo Wang. Segment anything in medical images. *arXiv preprint arXiv:2304.12306*, 2023. 8

[14] Yaron Meirovitch, Alexander Matveev, Hayk Saribekyan, David Budden, David Rolnick, Gergely Odor, Seymour Knowles-Barley, Thouis Raymond Jones, Hanspeter Pfister, Jeff William Lichtman, et al. A multi-pass approach to large-scale connectomics. *arXiv preprint arXiv:1612.02120*, 2016. 2

[15] Adam Michaleas, Lars A. Gjesteby, Michael Snyder, David Chavez, Meagan Ash, Matthew A. Melton, Damon G. Lamb, Sara N. Burke, Kevin J. Otto, Lee Kamentsky, Webster Guan, Kwanghun Chung, and Laura J. Brattain. Active learning pipeline for brain mapping in a high performance computing environment. In *2020 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–6, 2020. 8

[16] Diganta Misra, Trikay Nalamada, Ajay Uppili Arasanipalai, and Qibin Hou. Rotate to attend: Convolutional triplet attention module. *CoRR*, abs/2010.03045, 2020. 6

[17] Gábor Németh, Péter Kardos, and Kálmán Palágyi. Thinning combined with iteration-by-iteration smoothing for 3d binary images. *Graphical Models*, 73(6):335–345, 2011. 1, 2

[18] Hanchuan Peng, Michael Hawrylycz, Jane Roskams, Sean Hill, Nelson Spruston, Erik Meijering, and Giorgio A Ascoli. Bigneuron: large-scale 3d neuron reconstruction from optical microscopy images. *Neuron*, 87(2):252–256, 2015. 6

[19] Dylan Pollack, Lars A Gjesteby, Michael Snyder, David Chavez, Lee Kamentsky, Kwanghun Chung, and Laura J Brattain. Axon tracing and centerline detection using topologically-aware 3d u-nets. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 238–242. IEEE, 2022. 1, 2, 5, 6, 8

[20] Tobias Post, Christina Gillmann, Thomas Wischgoll, and Hans Hagen. Fast 3d thinning of medical image data based on local neighborhood lookups. In *EuroVis (Short Papers)*, pages 43–47, 2016. 1

[21] Valérie Robert, Yann Vasseur, and Vincent Brault. Comparing high-dimensional partitions with the co-clustering adjusted rand index. *Journal of Classification*, 38(1):158–186, 2021. 6

[22] Vanessa Robins. *Computational Topology for Point Data: Betti Numbers of α-Shapes*, pages 261–274. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002. 6

[23] Pierre Rougé, Nicolas Passat, and Odyssée Merveille. Cascaded multitask u-net using topological loss for vessel segmentation and centerline extraction. *arXiv preprint arXiv:2307.11603*, 2023. 1

[24] Nina I. Shamsi, Lars A. Gjesteby, David Chavez, Michael Snyder, Brian S. Eastwood, Matthew G. Fay, Nathan J. O'Connor, Jack R. Glaser, Charles R. Gerfen, and Laura J. Brattain. Self-supervised learning to improve topology-optimized axon segmentation and centerline detection. In *2023 IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2023. 3

[25] S. Shit, J. C. Paetzold, A. Sekuboyina, I. Ezhov, A. Unger, A. Zhylka, J. W. Pluim, U. Bauer, and B. 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, Los Alamitos, CA, USA, jun 2021. IEEE Computer Society. 1, 2, 3, 6

[26] N Steffensen and N Bourbakis. A method for parallel skeletonization of images. In *Proceedings., International Confer-*

ence on Image Processing, volume 2, pages 109–112. IEEE, 1995. 4

[27] Giles Tetteh, Velizar Efremov, Nils D. Forkert, Matthias Schneider, Jan Kirschke, Bruno Weber, Claus Zimmer, Marie Piraud, and Björn H. Menze. Deepvesselnet: Vessel segmentation, centerline prediction, and bifurcation detection in 3-d angiographic volumes. *Frontiers in Neuroscience*, 14, 2020. 1

[28] Yanshuai Tu, Duyan Ta, Zhong-Lin Lu, and Yalin Wang. Topology-preserving smoothing of retinotopic maps. *PLoS computational biology*, 17(8):e1009216, 2021. 2

[29] Martin G Wagner. Real-time thinning algorithms for 2d and 3d images using gpu processors. *Journal of real-time image processing*, 17(5):1255–1266, 2020. 1

[30] Qilong Wang, Banggu Wu, Pengfei Zhu, Peihua Li, Wangmeng Zuo, and Qinghua Hu. Eca-net: Efficient channel attention for deep convolutional neural networks. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pages 11531–11539. Computer Vision Foundation / IEEE, 2020. 6

[31] 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. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 3832–3841, 2020. 3, 4

[32] Madeleine K Wyburd, Nicola K Dinsdale, Ana IL Namburete, and Mark Jenkinson. Teds-net: enforcing diffeomorphisms in spatial transformers to guarantee topology preservation in segmentations. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 250–260. Springer, 2021. 2

[33] Wenjie Xie, Robert P Thompson, and Renato Perucchio. A topology-preserving parallel 3d thinning algorithm for extracting the curve skeleton. *Pattern Recognition*, 36(7):1529–1544, 2003. 1, 4

[34] Minghui Zhang, Guang-Zhong Yang, and Yun Gu. Differentiable topology-preserved distance transform for pulmonary airway segmentation. *arXiv preprint arXiv:2209.08355*, 2022. 1, 2

[35] Kelly H Zou, Simon K Warfield, Aditya Bharatha, Clare MC Tempany, Michael R Kaus, Steven J Haker, William M Wells III, Ferenc A Jolesz, and Ron Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic radiology*, 11(2):178–189, 2004. 6