

This CVPR 2020 workshop paper is the Open Access version, provided by the Computer Vision Foundation. Except for this watermark, it is identical to the accepted version; the final published version of the proceedings is available on IEEE Xplore.

A topological encoding convolutional neural network for segmentation of 3D multiphoton images of brain vasculature using persistent homology

Mohammad Haft-Javaherian^{1,2*} Martin Villiger¹ Chris B. Schaffer³ Nozomi Nishimura³ Polina Golland² Brett E. Bouma^{1,4}

 ¹ Wellman Center for Photomedicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114 USA
 ² Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts Institute of Technology, Cambridge, MA 02142 USA

 ³ Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY 14853 USA
 ⁴ Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02142 USA

{haft,polina}@csail.mit.edu {nn62,cs385}@cornell.edu
 {mvilliger, bouma}@mgh.harvard.edu

Abstract

The clinical evidence suggests that cognitive disorders are associated with vasculature dysfunction and decreased blood flow in the brain. Hence, a functional understanding of the linkage between brain functionality and the vascular network is essential. However, methods to systematically and quantitatively describe and compare structures as complex as brain blood vessels are lacking. 3D imaging modalities such as multiphoton microscopy enables researchers to capture the network of brain vasculature with high spatial resolutions. Nonetheless, image processing and inference are some of the bottlenecks for biomedical research involving imaging, and any advancement in this area impacts many research groups. Here, we propose a topological encoding convolutional neural network based on persistent homology to segment 3D multiphoton images of brain vasculature. We demonstrate that our model outperforms state-of-the-art models in terms of the Dice coefficient and it is comparable in terms of other metrics such as sensitivity. Additionally, the topological characteristics of our model's segmentation results mimic manual ground truth. Our code and model are open source at https://github.com/mhaft/DeepVess.

The health of the brain and heart, as the most two critical organs, interconnect through the brain vasculature. Since the brain energy depends exclusively on the supply of oxygen and glucose through the bloodstream and the brain energy reserve is very limited, the brain requires steady and reliable blood perfusion [24]. Therefore, any blood flow interruption causes temporary or permanent functional impairments [38]. Researchers and physicians utilize various imaging modalities, such as multiphoton microscopy (MPM) [4] and optical coherence tomography [39], to study and examine the 3D geometrical, topological, and fluid dynamics characteristics of the brain vasculature, including capillaries in live animal models.

Accurate analysis of brain vasculature over large 3D volumes and multiple subjects within several controlled groups forms a considerable portion of the bottleneck processes in terms of time and personal allocation in biomedical research and medicine. For instance, researchers studied the effects of brain capillary stalling on Alzheimer's disease in animal models and reported that the image analysis required order of magnitude higher time allocations in comparison to the actual image acquisitions [6, 5, 11, 13, 20]. Notably, the vessel segmentation and vectorization are the prevalent required image analysis tasks [26]. Researchers often consider the segmentation as a classification task and have pro-

^{1.} Introduction

^{*}Corresponding author.



Figure 1. An example of a persistence diagram. Panels A and B are the target and query images, respectively. The image intensities are normalized into the [0, 1] range, and the color blue, red, and dark yellow represent low, medium, and high intensity values. In panel C, green and magenta markers are associated with the target and query images, while square and circle markers are related to H_0 and H_1 homology groups, respectively. The unfilled green markers located on the horizontal axis are the projected elements to match the cardinalities of the two persistence diagrams. The arrows in panel C are the distance between the two persistence diagrams.

posed supervised, unsupervised, and semi-supervised solutions using classical computer vision methods [26] and recently deep learning (DL) methods [27].

In this work, we present a novel method for the segmentation of 3D in vivo MPM images of brain vasculature based on a topological encoding convolutional neural network (CNN) using persistent homology (PH) as an unsupervised loss term in addition to two other supervised loss terms. Our work is motivated by the fact that the-state-ofthe-art MPM segmentation methods such as DeepVess [21] suffer from topological errors in the regions with low signalto-noise (SNR) and by the recent successful integration of the algebraic topology metrics such as persistent homology into differentiable DL layers [7]. Even though persistent homology has been introduced and utilized in various fields such as computational topology, its application in the computer vision is limited to preprocessed feature selections [23]. Our algorithm improved DeepVess' accuracy significantly, specifically in the low SNR regions, and by producing more plausible results, qualitatively.

2. Related Work

2.1. Vessel segmentation

The segmentation of vasculature networks or particular vessel segments, *e.g.* the aorta, is an essential task in most biomedical research applications and medical settings due to the fact that the original task requires enormous time and personal allocations. The vessel segmentation problem is tackled through the established computer vision methods, which commonly require feature engineering and accept/reject rules, as well as the DL methods, which are formulated as supervised, unsupervised, or semi-supervised learning problems [14, 29, 42]. For instance, Yi *et al.* [45] proposed a 3D method using the region growing algorithm

utilizing its computational and implementation efficiency. Similarly, Mile *et al.* [29] represented the segmentation in terms of a deformable parametric model that accurately match the vessel fractal shapes.

On the other hand, due to the current success of DL in various research fields, DL models are popular and yield higher accuracy levels. In recent years, researchers, practitioners, and challenge participants almost exclusively utilize different variants of DL models to tackle different biomedical computer vision problems. For example, a model based on the transfer learning of the Google Inception V3 network [40] surpassed the accuracy of seven ophthalmologists combined in the diabetic retinopathy detection task within the color fundus images [18]. The CNN models, such as the model by Wu et al. [44] can segment vasculature networks successfully, and in different circumstances, recurrent neural network (RNN) models are proposed in place of post-processing conditional random fields (CRF) models [16, 15]. The most common biomedical image segmentation DL architecture theme is called U-Net [35], which is a fully convolutional architecture [28] with encoding and decoding layers in addition to skip connections at different spatial scales.

2.2. MPM segmentation

MPM is the standard imaging modality for studying different live animal models enabling longitudinal studies (*e.g.* weeks and months) while providing high temporal resolutions (*e.g.* 2-30 Hz) and sub-micron spatial resolutions with minimally invasive procedure requirements. In comparison to the linear process-based counterpart imaging modality, the intensity of the non-linear processes in MPM imaging (*e.g.* two-photon excited fluorescence) is proportional to the square or cube of the required average laser power. This non-linearity increases imaging depth and decrease photo-

bleaching [25].

Tsai *et al.* [42] introduced the first comprehensive MPM image processing software suit, a.k.a. VIDA, based on traditional computer vision methods with the ability to handle large datasets. On the other hand, Teikari *et al.* [41] utilized 2.5D CNN architecture with CRF post-processing to segment MPM vessel images, while Bates *et al.* [3] used RNN architecture to extract the vasculature graph representations. Recently, DeepVess *et al.* [21] was proposed, which is an optimized CNN architecture with a customized loss function and accuracy comparable to that of expert human annotators for both vessel segmentation and centerline extractions and Gur *et al.* [19] developed an unsupervised neural network with active contours mimicking loss term.

2.3. Topological data analysis

Topological data analysis (TDA) is a mathematical field that studies complex datasets from the algebraic topology point of view. PH, as a TDA theory, is able to characterize high-dimensional manifolds within datasets such as graphbased characterization, unsupervised feature engineering, and image manifold analysis [12, 33, 36].

In comparison to other models with no regularization or preset boundary topologies (*e.g.* multivariate regression model and support vector machines), Chen *et al.* [8] used PH to regularize the decision boundaries and to simplify the boundaries. Similarly, Rieck *et al.* [34] used the PH as a complexity metric for DL architectures by modeling the network as a weighted stratified graph to facilitate the underling optimization problem required for the learning process such as the mimicking the best DL regularization tricks (*e.g.* dropout and batch normalization) and optimization practice such as validation-based early stopping.

The feature engineering based on the input image is the primary utilization of PH for the segmentation task. Qaiser et al. [32] trained a CNN based on the extracted topological features using enhanced persistent homology profiles to segment tumors in Hematoxylin and Eosin stained histology images. Differently, Assaf et al. [2] segmented the cells by selecting the target object within the candidates based on the largest life span using PH of the cubical complex. Similarly, Gao et al. [17] used the PH to segment the high-resolution CT images of the left ventricle including the papillary muscles and the trabeculae by detecting the salient object with the desirable topological characteristic, *i.e.* handles, to be restored in the final results. Furthermore, Clough et al. [10] tackled the same problem with the exception of using the PH for providing the gradient signals to the learning optimizer.

3. Method

3.1. Persistent homology

The homology in the context of abstract algebra uses invariant theory to maps the groups on an algebraic variety such as Euclidean spaces to the topological space. If we describe the manifolds in a D-dimensional Euclidean space $\mathcal{X} \in \mathbb{R}^D$ as a simplicial complex \mathcal{K} , which is a generalization of graphs in the higher dimensions, simplicial homology represents the connectivity of \mathcal{K} in terms of homology groups using matrix reduction algorithms. A k-dimensional simplex is the convex combination of k + 1 vertices. The vertices correspond to points in a sufficiently nice ambient \mathbb{R}^D space, which results in the existence of the simplex geometrical realization within \mathcal{X} . A collection of simplices forms a simplicial complex \mathcal{X} if for every simplex $\tau \in \mathcal{X}$, every face $\sigma \subseteq \tau$ is also in \mathcal{X} and also for any two simplices τ_i and τ_i , $\sigma = \tau_i \cap \tau_j$ is a face of both τ_i and τ_j .

Homology groups $\{H_k(\mathcal{K}) : k \in \{0, ..., D-1\}\}$ describe the k^{th} order topological features, *i.e.* holes of the k^{th} dimension. For instance, $H_0(\mathcal{K})$, $H_1(\mathcal{K})$, and $H_2(\mathcal{K})$ correspond to connected components, 2D hole or 3D tunnel, and 3D voids, respectively. The *Betti numbers* (β_i) summarize the homology group information by measuring the cardinality of the $H_i(\mathcal{K})$. For example, a 2D ring has Betti numbers of (1, 1), a filled 2D square has Betti numbers of (1, 0), a torus has Betti numbers of (1, 1, 1), and a sphere has Betti numbers of (1, 0, 1).

Betti numbers are very abstract representation of the homology groups and they are not smooth in noisy datasets. Persistent homology resolves this shortcoming of Betti numbers by measuring the topological features of \mathcal{X} at different resolutions to promote the detection of the true features embedded in \mathcal{X} and exclusion of features constructed based on quantification and noise errors. The simplicial complexes can be ordered into a nested sequence by considering a non-decreasing function $f : \mathcal{K} \to \mathbb{R}$ that induces a homomorphism on the simplicial homology groups for each dimension. For every $a \in \mathbb{R}$, there is a subcomplex of \mathcal{K} as a sublevel set $\mathcal{K}(a) = f^{-1}(-\inf, a]$ and a set of scale weights $a_0 \leq a_1 \leq ... \leq a_n$ (a.k.a. filtration parameter) put the \mathcal{K} in filtration based on the ordering of f to represent the growth of \mathcal{K} by changing the filtration parameter (a_i) ,

$$\emptyset = \mathcal{K}_0 \subseteq \mathcal{K}_1 \subseteq \dots \subseteq \mathcal{K}_n = \mathcal{K}.$$
 (1)

Over the course of \mathcal{K} growth by filtration, topological features are born and die at different scales. Persistence homology maps the birth and death of features into \mathbb{R}^2 space as (b_i, d_i) points, where $b_i, d_i \in [0, 1]$ or any other range. The collection of all points based on k-dimensional topological features generate the k^{th} persistence diagram,

$$\mathcal{PD}_k: (\mathcal{X}, f) \to \{(b_i, d_i) : i \in \mathcal{I}_k\}.$$
 (2)



Figure 2. Our convolutional neural network architecture. The input is a $98 \times 98 \times 7$ 3D image patch with a single input channel. The first three 3D convolution layers with the $3 \times 3 \times 3$ kernel size and 32 features were followed by a max-pooling with the 2×2 kernel size and 2×2 strides. Then two 2D convolution layers with the 3×3 kernel size and 64 features followed by a max-pooling with the 2×2 kernel size and 2×2 strides are applied. Finally, a dense layer classifier with a 1024-node fully-connected hidden layer as well as a dropout layer at the training time was applied. The loss function terms are based on the generalize Dice, cross-entropy, and topological persistence.

If homology is computed over a field, the homology of a filtration admits an interval decomposition, which is the sum of the elements of rank one within the interval $[b_i, d_i)$. The persistence of each point in each \mathcal{PD}_k is the lifetime of the feature and defined as $pers(b_i, d_i) = |d_i - b_i|$. Furthermore, the points in each \mathcal{PD}_k are indexed in a descending order based on their persistence value. Therefore, the first point in \mathcal{I}_k is the most persistent feature and the last one is probably an artifact.

To implement the PH as a differentiable DL layer, the complete inverse map from \mathcal{PD}_k to the input dataset \mathcal{D} : (X, Y) and a loss function are required. The complete inverse map has two steps: mapping \mathcal{PD}_k to \mathcal{X} and then mapping \mathcal{X} to \mathcal{D} . The inverse map between \mathcal{PD}_k and \mathcal{X} can be defined by searching for the pair of two simplices τ and τ' such that $b_i = f(\tau)$ and $d_i = f(\tau')$,

$$\pi(\mathcal{PD}_k, i \in \mathcal{I}_k; f) : (b_i, d_i) \to (\tau, \tau').$$
(3)

The second step of the complete inverse map is to connect the \mathcal{X} to the input dataset \mathcal{D} . Since in most computer vision problems, the $\mathcal{D} = (X, Y)$ consists of 3D digitized images as voxel grids, we define the filtration function on the voxel intensity value and simplex based on the voxel neighborhood. Therefore, the filtration is equivalent of level sets at different intensity thresholds and the mapping associates each simplex (τ) to the indices of voxels $\sigma \in \tau$ that their intensity values are equal to the filtering cutoff value (a_i) ,

$$\omega(\tau) = \operatorname{argmax}_{\sigma \in \tau} f(\sigma). \tag{4}$$

The PH cost functions should represent the distance between the target persistence diagram \mathcal{PD} and the predicted persistence diagram \mathcal{PD} . We use *Wasserstein-1 distance* $\mathcal{W}_1(p,q)$, a.k.a. Earth-Mover distance, which is related to the minimum cost of stockpile transportation from stock configuration p to stock configuration q, where the transportation cost is equal to the mass times the transportation distance. Note that under a few mild assumptions, \mathcal{W}_1 is continuous and differentiable almost everywhere [31]. The objective is to minimize the \mathcal{W}_1 between \mathcal{PD} and \mathcal{PD} . At first, the cardinalities of \mathcal{PD} and \mathcal{PD} are matched by projecting the additional points in \mathcal{PD} to the diagonal (5) and then the loss function measures the Wasserstein-1 distance



Figure 3. 2D slice-based Dice coefficient variation as a function of imaging depth based on our model and baselines. On each box, the central red line indicates the median, while the bottom and top limits of the box indicate the first and third quartiles, respectively. The whiskers do not include outliers, which are plotted using the plus sign.

between counterpart topological features to equate the number of topological features with high persistent between \mathcal{PD} and $\hat{\mathcal{PD}}$ and push the rest of that features to have zero persistent values,

$$proj(b_{i}, d_{i}) = \frac{b_{i} + d_{i}}{2},$$

$$\mathcal{E}_{\mathcal{T}}(\beta_{k}; \hat{\mathcal{PD}}_{k}) = \left(\sum_{i=1}^{\beta_{k}} \|(b_{i}, d_{i}) - (0, 1)\|^{1}\right)^{1} + \left(\sum_{i=\beta_{k}+1}^{|\mathcal{I}_{k}|} \|(b_{i}, d_{i}) - proj(b_{i}, d_{i})\|^{1}\right)^{1} = \sum_{i=1}^{n_{k}} |1 - d_{i} + b_{i}| + \sum_{i=n_{k}+1}^{|\mathcal{I}_{k}|} |d_{i} - b_{i}|.$$
(6)

The gradient of the topological loss function (6) with respect to the input image is measurable through the inverse maps (3) and enables the optimization of the whole network in an end-to-end fashion.

An example of this framework is depicted in Figure 1. Panel A and B in Figure 1 are the target and query images, respectively, and the image intensities are normalized into the [0, 1] range. The color blue, red, and dark yellow represent low, medium, and high intensity values such as 0.01, 0.20, and 0.70. In Figure 1.C, green and magenta markers are associated with the target and query images, and square and circle markers are related to H_0 and H_1 homology groups, respectively. The unfilled markers located on the horizontal axis are the elements that are projected to the diagonal line to match the cardinalities of the target persistence diagram \mathcal{PD} and the query persistence diagram \mathcal{PD} . The arrows are the distances between the target and query features, which measure the distance between the two persistence diagrams. For example, when the filtering value is equal to the red color (e.g. 0.20), the letter "j" will appear, which results in the birth of a hole that is confined between the letter "C" and the letter "j" in addition to birth of a new connected components that corresponds to the dot of letter "j". The changes due to appearance of letter "j" are reflected in Figure 1.C as the magenta circle and square with mid-range persistent values. The magenta circle and square with low-range persistent values in the lower right corner are related to the letter "Q".

3.2. Network architecture

We adopted the network architecture of DeepVess [21], which is the state-of-the-art architecture optimized for vessel segmentation in MPM images, as our baseline architecture (Fig. 2). Briefly, DeepVess' architecture has an input image, which is 7-voxel deep along the third direction and initially apply three 3D convolution layers with the $3 \times 3 \times 3$ kernel size, 32 features, and without padded edges followed by a max-pooling with the 2×2 kernel size and 2×2 strides. Then two 2D convolution layers with the 3×3 kernel size and 64 features followed by a max pooling with the 2×2 kernel size and 2×2 strides are applied. Finally, the output of the last max-pooling layer is flattened to a fullyconnected layer followed by a 1024-node fully-connected layer in addition to a dropout layer at the training time (with 50% probability) and the last fully-connected layer, which is reshaped to the output patch size. LeakyReLU was used as the activation function for all the layers. The width of the input image and consequently the receptive field as the output size were the optimized design hyperparameters that were selected based on the new loss function and prior knowledge of the topological characteristics of brain vasculature networks.

3.3. Learning and implementation

We preprocessed the 3D *in vivo* MPM images of mouse brain vasculature networks following [21]. Briefly, the physiological motion artifacts were removed using nonrigid non-parametric diffeomorphic demons image registration [43], followed by image intensity normalization and $1\mu m^3$ spatial resolution scaling.

DeepVess resolved the highly unbalanced class population problem by using a customized cross-entropy loss function (7),

$$\mathcal{L}_{DeepVess}(y,\hat{y}) = \sum_{i \in \{TP, FP, FN\}} \mathcal{CE}(y_i, \hat{y}_i), \quad (7)$$

which is measured over all voxels except true negative voxels. Instead, we used generalized Dice loss function (8) due to the successful results in [30],

$$\mathcal{L}_{Dice}(y,\hat{y}) = 1 - \frac{2\sum_{i}(y_i \times \hat{y}_i) + \epsilon}{\sum_{i}(y_i^2 + \hat{y}_i^2) + \epsilon}.$$
(8)

Additionally, we used the normal cross-entropy loss function (\mathcal{L}_{CE}) for the initial optimization warm-up to decrease the required training time,

$$C\mathcal{E}(y_i, \hat{y}_i) = -y_i \times log(\hat{y}_i) - (1 - y_i) \times log(1 - \hat{y}_i).$$
(9)

In this training regime, initially, \mathcal{L}_{CE} is the significant loss term, but after the initial warm-up phase, $\mathcal{L}_{DeepVess}$ will automatically dominate the optimization as a result of the highly unbalanced class populations.

The objective of the topological cost function (6) is to enforce the prior knowledge of the desired topological characterises in terms of the target Betti numbers (β_k). Due to the fractal characteristics of venous and arterial networks and their interaction through the grid networks of capillaries [37]. The topological characterises of brain vasculature network is scale-dependent. Therefore, the optimal size of the field of view based on the topological constrains is the largest field of view, which has similar Betti numbers in various regions. In our case, which is the mouse brain vasculature network, we observed there is no large capillary loop within a $21 \times 21 \ \mu m^2$ region, which translates to target β_0 of 1 and β_1 of 0, equivalent to one connected component without any hole. Therefore, the topological loss function is deterministic without the need to be inferred from the data, and it can be formulated based on (6) into

$$\mathcal{L}_{\mathcal{T}}(\hat{\mathcal{P}}\mathcal{D}) = \sum_{k=0}^{D-1} \mathcal{E}_{\mathcal{T}}(\tilde{\beta}_k; \hat{\mathcal{P}}\mathcal{D}_k)$$

$$= \mathcal{E}_{\mathcal{T}}(1; \hat{\mathcal{P}}\mathcal{D}_0) + \mathcal{E}_{\mathcal{T}}(0; \hat{\mathcal{P}}\mathcal{D}_1).$$
(10)

The final proposed loss function contains three terms: Dice, cross-entropy, and topology loss functions. We measured the topological loss function on 10% of the samples in the mini-batch to reduce the computational cost, while maintaining a sufficient gradient signal,

$$\mathcal{L}(y,\hat{y}) = \mathcal{L}_{Dice}(y,\hat{y}) + \lambda \mathcal{L}_{CE}(y,\hat{y}) + \lambda' \mathcal{L}_{\mathcal{T}} \left(\mathcal{PD}(\hat{y}_1,..,\hat{y}_{\lfloor n/10 \rfloor}) \right).$$
(11)

We implemented our model in TensorFlow [1] and measure PH using the "*Topology Layer*" package [7]. We trained our model using *Adam* stochastic optimization with a learning rate of 10^{-4} for 2000 epochs. We set the λ and λ' to the optimal values of 1 and 0.01 based on the evaluation results of the validation dataset. The division patterns of the image patches within the training dataset were shifted randomly to cover different possible slicing schemes. Also, the same procedure was used to enhance the classification at the inference phase. Our code and model are open source at *https://github.com/mhaft/DeepVess*.

4. Experiments

4.1. Data

We demonstrate the performance of our algorithm on DeepVess dataset [21], which is the largest, publicly available, manually-labeled, *in vivo* MPM image dataset of mouse brain vasculature with 24 mice and more than $13 \times 10^6 \mu m^3$ labelled images. We utilized their division of the dataset into training, validation, and testing sub-datasets (50%, 25%, and 25%). Similarly, while we fine-tuned and optimized the models based on the performance on the validation dataset, the final model accuracy results were measured on the independent test sub-dataset.

4.2. Baseline and evaluation

We compare our proposed model to the state-of-the-art segmentation methods VesselNN [41], DeepVess [21], and UMIS [19] in addition to VIDA [42] and 3D U-Net [9], which are the methods commonly used in the analysis of MPM and other biomedical image modalities.

We evaluate our model qualitatively with figures in addition to quantitatively based on *accuracy*, *sensitivity*, *specificity*, *Dice coefficient* (a.k.a. *F1*), and *Jaccard index*. This metrics are based on the populations of *true positive* (*TP*), *true negative* (*TN*), *false positive* (*FP*), and *false negative* (*FN*) cases,

$$Sensitivity = \frac{TP}{TP + FN},$$
(12)

$$Specificity = \frac{TN}{TN + FP},\tag{13}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN},$$
 (14)

$$Jaccard = \frac{IP}{TP + FP + FN},$$
(15)

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN}.$$
 (16)

4.3. Results and Discussion

We constructed a loss function (11) that enables the optimization of CNN models to segment a highly imbalanced foreground using the Dice term, besides encoding the prior knowledge of the topological characteristics of the segmentation results into the dense layer classifier through the topological loss term.

We compared the performance of our proposed model to the-state-of-the-art baselines quantitatively in Table 1 and



Figure 4. Qualitative comparison between our model and DeepVess [21]. A-B. The cross-sections spaced through the depth are plotted. The first column is the image intensity of the 2D slice in gray-scale. The second and third columns are the results of DeepVess and our models overlaid on the intensity images. The last column of each panel is the entropy with cooler and warmer colors representing the lower and higher uncertainty levels, respectively. C. The results of DeepVess and our models applied to the last sample. The cyan and yellow arrows represent the discontinuity and hole artifacts, respectively. Scale bar is $100\mu m$.

Model	Accuracy	Sensitivity	Specificity	Dice	Jaccard
VIDA [42, 19]	60.2%	99.8%	56.4%	30.4%	17.9%
VesselNN [41, 21]	95.1%	62.4%	98.7%	69.7%	55.1%
3D U-Net [9, 22]	95.6%	70.0%	98.2%	72.7%	59.4%
UMIS [19]	98.6%	99.2%	98.6%	82.9%	70.8%
DeepVess [21]	96.4%	90.0%	97.0%	81.6%	69.2%
Ours	96.8%	96.4%	98.2%	86.5%	76.2%

Table 1. Result of our model in comparison to the baseline models. Baseline models are VIDA [42], VesselNN [41], U-Net [9], UMIS [19], and DeepVess [21]. The measured performance metrics are accuracy, sensitivity, specificity, Dice coefficient, and Jaccard index, with the last two metrics being correlated. Our model presents a higher Dice and Jaccard indices while maintaining comparable performance in terms of the other metrics.

qualitatively in Figure 4. Our model outperforms other models in terms of Dice and Jaccard indices. Since foreground voxels are less than 10% of the population, these two metrics are better indicators of the actual performance of a model in comparison to other metrics such as accuracy. A classifier that classifies all the input voxels as backgrounds results in 90% accuracy and 100% specificity (or 100% sensitivity if it classifies all the voxels as foregrounds); nevertheless, its Dice and Jaccard indices are 0%. Under other models, the results will have slightly higher accuracy, sensitivity, and specificity; but with lower Dice and Jaccard indices in addition to having topological errors (Figure 4). However, these shortcomings demonstrate the existing capacity for improvement, mainly because this tool is frequently used in different branches of biomedical research and also the dependency of downstream image analysis tasks, such as the centerline extraction, on the quality of the segmentation results.

The usage of infrared light to excite fluorophores with high absorption cross-sections enables deeper imaging in light-scattering tissues with low energy requirements and high resolution. However, the main sources of image processing difficulties are the rapid loss of signal and the background signal from the superficial layer in the images acquired deeper within the tissue. Due to this intrinsic limitation, the study of the model performance as a function of depth is essential. Figure 3 is the box-plot of the Dice Coefficient measured at each slice for all slices in the 3D image. Our model has lower Dice variation in comparison to DeepVess, while our model maintains a similar median value for the performance, which is an indicator that our model is less sensitive to the SNR variation over the depth of imaging with higher consistency.

Figure 4 illustrates the results of our model and Deep-Vess overlaid on the input images. In Figure 4.A-B, the cross-sections spaced through the depth are plotted. The first column is the image intensity of the 2D slice in grayscale. The second and third columns are the results of Deep-Vess and our model overlaid on the intensity images. The input images were divided into the input patches using different splitting schemes to measure our model uncertainty by measuring Shannon's entropy. Higher entropy values correlate with higher segmentation uncertainties. The last column is the entropy with cooler and warmer colors representing the lower and higher uncertainty levels, respectively.

Although the deeper slices, e.g. Figure 4.B, have more voxels with uncertainty; evidently, the level of uncertainty is lower than the results in [21] and the topological features of the segmentation results are close to the ground truth and the topological priors in our application. In Figure 4.C, the cyan and yellow arrows in DeepVess' result represent the discontinuity and hole artifacts, respectively. The segmentation results based on our model do not have any hole in addition to the fact that the integrity and continuity of the vessel are very well-maintained. These two characteristics are essential for the downstream image analysis tasks, including the vessel centerline extraction.

5. Conclusion

We proposed a convolutional neural network architecture with a loss function consists of three terms, including a topological loss term. The topological loss term is based on the persistent homology and encodes the prior knowledge of topological features and their characteristics into the model. We demonstrated the application of this model to the segmentation of brain vasculature networks in multiphoton microscopy images in mice. While our model outperforms the-state-of-the-art models based on the Dice and Jaccard indices, our model has comparable accuracy, sensitivity, and specificity in addition to producing segmentation results with less topological errors in comparison to thestate-of-the-art models.

Acknowledgment

This work was supported by the National Institutes of Health NIBIB under Grant P41EB-015903 and Grant P41EB015902. In addition, MH was supported by Bullock Postdoctoral Fellowship.

References

- Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467, 2016.
- [2] Rabih Assaf, Alban Goupil, Mohammad Kacim, and Valeriu Vrabie. Topological persistence based on pixels for object segmentation in biomedical images. In 2017 Fourth International Conference on Advances in Biomedical Engineering (ICABME), pages 1–4. IEEE, 2017.
- [3] Russell Bates, Benjamin Irving, Bostjan Markelc, Jakob Kaeppler, Ruth Muschel, Vicente Grau, and Julia A Schnabel. Extracting 3d vascular structures from microscopy images using convolutional recurrent networks. *arXiv preprint arXiv:1705.09597*, 2017.
- [4] Pablo Blinder, Philbert S Tsai, John P Kaufhold, Per M Knutsen, Harry Suhl, and David Kleinfeld. The cortical angiome: an interconnected vascular network with noncolumnar patterns of blood flow. *Nature neuroscience*, 16(7):889, 2013.
- [5] Oliver Bracko, Brendah N Njiru, Madisen Swallow, Muhammad Ali, Mohammad Haft-Javaherian, and Chris B Schaffer. Increasing cerebral blood flow improves cognition into late stages in alzheimer's disease mice. *Journal of Cerebral Blood Flow & Metabolism*, page 0271678X19873658, 2019.
- [6] OLIVER Bracko, Lindsay K Vinarcsik, Jean C Cruz Hernandez, Nancy E Ruiz-Uribe, Mohammad Haft-Javaherian, Kaja Falkenhain, Egle M Ramanauskaite, Muhammad Ali, Aditi Mohapatra, Madisen A Swallow, et al. High fat diet worsens pathology and impairment in an alzheimers mouse model, but not by synergistically decreasing cerebral blood flow. *bioRxiv*, 2019.
- [7] Rickard Brüel-Gabrielsson, Bradley J Nelson, Anjan Dwaraknath, Primoz Skraba, Leonidas J Guibas, and Gunnar Carlsson. A topology layer for machine learning. arXiv preprint arXiv:1905.12200, 2019.
- [8] Chao Chen, Xiuyan Ni, Qinxun Bai, and Yusu Wang. A topological regularizer for classifiers via persistent homology. arXiv preprint arXiv:1806.10714, 2018.
- [9] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.
- [10] James R Clough, Ilkay Oksuz, Nicholas Byrne, Veronika A Zimmer, Julia A Schnabel, and Andrew P King. A topological loss function for deep-learning based image segmentation using persistent homology. arXiv preprint arXiv:1910.01877, 2019.
- [11] Jean C Cruz Cruz Hernández, Oliver Bracko, Calvin J Kersbergen, Victorine Muse, Mohammad Haft-Javaherian, Maxime Berg, Laibaik Park, Lindsay K Vinarcsik, Iryna Ivasyk, Daniel A Rivera, et al. Neutrophil adhesion in brain capillaries reduces cortical blood flow and impairs memory

function in alzheimer's disease mouse models. *Nature neuroscience*, 22(3):413–420, 2019.

- [12] Herbert Edelsbrunner and John Harer. *Computational topology: an introduction*. American Mathematical Soc., 2010.
- [13] Kaja Falkenhain, Nancy E Ruiz-Uribe, Mohammad Haft-Javaherian, Muhammad Ali, Pietro E Michelucci, Chris B Schaffer, OLIVER Bracko, et al. Voluntary running does not increase capillary blood flow but promotes neurogenesis and short-term memory in the app/ps1 mouse model of alzheimers disease. *bioRxiv*, 2020.
- [14] Muhammad Moazam Fraz, Paolo Remagnino, Andreas Hoppe, Bunyarit Uyyanonvara, Alicja R Rudnicka, Christopher G Owen, and Sarah A Barman. Blood vessel segmentation methodologies in retinal images–a survey. *Computer methods and programs in biomedicine*, 108(1):407– 433, 2012.
- [15] Huazhu Fu, Yanwu Xu, Stephen Lin, Damon Wing Kee Wong, and Jiang Liu. Deepvessel: Retinal vessel segmentation via deep learning and conditional random field. In *International conference on medical image computing and computer-assisted intervention*, pages 132–139. Springer, 2016.
- [16] Huazhu Fu, Yanwu Xu, Damon Wing Kee Wong, and Jiang Liu. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In 2016 IEEE 13th international symposium on biomedical imaging (ISBI), pages 698–701. IEEE, 2016.
- [17] Mingchen Gao, Chao Chen, Shaoting Zhang, Zhen Qian, Dimitris Metaxas, and Leon Axel. Segmenting the papillary muscles and the trabeculae from high resolution cardiac et through restoration of topological handles. In *International Conference on Information Processing in Medical Imaging*, pages 184–195. Springer, 2013.
- [18] Varun Gulshan, Lily Peng, Marc Coram, Martin C Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, Kasumi Widner, Tom Madams, Jorge Cuadros, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*, 316(22):2402–2410, 2016.
- [19] Shir Gur, Lior Wolf, Lior Golgher, and Pablo Blinder. Unsupervised microvascular image segmentation using an active contours mimicking neural network. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 10722–10731, 2019.
- [20] Mohammad Haft Javaherian. Quantitative assessment of cerebral microvasculature using machine learning and network analysis. 2019.
- [21] Mohammad Haft-Javaherian, Linjing Fang, Victorine Muse, Chris B Schaffer, Nozomi Nishimura, and Mert R Sabuncu. Deep convolutional neural networks for segmenting 3d in vivo multiphoton images of vasculature in alzheimer disease mouse models. *PloS one*, 14(3), 2019.
- [22] Mohammad Haft-Javaherian, Chris B Schaffer, Nozomi Nishimura, and Mert R Sabuncu. Comparison of convolutional neural and fully convolutional networks for segmentation of 3d in vivo multiphoton microscopy images of brain vasculature. In *Optics and the Brain*, pages BT2A–4. Optical Society of America, 2019.

- [23] Christoph Hofer, Roland Kwitt, Marc Niethammer, and Andreas Uhl. Deep learning with topological signatures. In *Advances in Neural Information Processing Systems*, pages 1634–1644, 2017.
- [24] KA Hossmann. Viability thresholds and the penumbra of focal ischemia. *Annals of neurology*, 36(4):557–565, 1994.
- [25] Adam M Larson. Multiphoton microscopy. *Nature Photonics*, 5(1):1, 2010.
- [26] David Lesage, Elsa D Angelini, Isabelle Bloch, and Gareth Funka-Lea. A review of 3d vessel lumen segmentation techniques: Models, features and extraction schemes. *Medical image analysis*, 13(6):819–845, 2009.
- [27] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical image analysis*, 42:60–88, 2017.
- [28] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.
- [29] Julien Mille and Laurent D Cohen. Deformable tree models for 2d and 3d branching structures extraction. In 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pages 149–156. IEEE, 2009.
- [30] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In 2016 Fourth International Conference on 3D Vision (3DV), pages 565–571. IEEE, 2016.
- [31] Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport. *Foundations and Trends* (®) *in Machine Learning*, 11(5-6):355–607, 2019.
- [32] Talha Qaiser, Yee-Wah Tsang, David Epstein, and Nasir Rajpoot. Tumor segmentation in whole slide images using persistent homology and deep convolutional features. In Annual Conference on Medical Image Understanding and Analysis, pages 320–329. Springer, 2017.
- [33] Bastian Rieck, Ulderico Fugacci, Jonas Lukasczyk, and Heike Leitte. Clique community persistence: A topological visual analysis approach for complex networks. *IEEE transactions on visualization and computer graphics*, 24(1):822– 831, 2017.
- [34] Bastian Rieck, Matteo Togninalli, Christian Bock, Michael Moor, Max Horn, Thomas Gumbsch, and Karsten Borgwardt. Neural persistence: A complexity measure for deep neural networks using algebraic topology. In *International Conference on Learning Representations* (ICLR), 2019.
- [35] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. Unet: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [36] Ann Sizemore, Chad Giusti, and Danielle S Bassett. Classification of weighted networks through mesoscale homological features. *Journal of Complex Networks*, 5(2):245–273, 2017.

- [37] Amy F Smith, Vincent Doyeux, Maxime Berg, Myriam Peyrounette, Mohammad Haft-Javaherian, Anne-Edith Larue, John H Slater, Frédéric Lauwers, Pablo Blinder, Philbert Tsai, et al. Brain capillary networks across species: a few simple organizational requirements are sufficient to reproduce both structure and function. *Frontiers in physiology*, 10:233, 2019.
- [38] William E Sonntag, Delrae M Eckman, Jeremy Ingraham, and David R Riddle. Regulation of cerebrovascular aging. In David R Riddle, editor, *Brain aging: models, methods, and mechanisms*, pages 279–304. CRC Press/Taylor & Francis, Boca Raton, FL, 2007.
- [39] Vivek J Srinivasan, Harsha Radhakrishnan, Eng H Lo, Emiri T Mandeville, James Y Jiang, Scott Barry, and Alex E Cable. Oct methods for capillary velocimetry. *Biomedical* optics express, 3(3):612–629, 2012.
- [40] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.
- [41] Petteri Teikari, Marc Santos, Charissa Poon, and Kullervo Hynynen. Deep learning convolutional networks for multiphoton microscopy vasculature segmentation. arXiv preprint arXiv:1606.02382, 2016.
- [42] Philbert S Tsai, John P Kaufhold, Pablo Blinder, Beth Friedman, Patrick J Drew, Harvey J Karten, Patrick D Lyden, and David Kleinfeld. Correlations of neuronal and microvascular densities in murine cortex revealed by direct counting and colocalization of nuclei and vessels. *Journal of Neuroscience*, 29(46):14553–14570, 2009.
- [43] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, and Nicholas Ayache. Diffeomorphic demons: Efficient nonparametric image registration. *NeuroImage*, 45(1):S61–S72, 2009.
- [44] Aaron Wu, Ziyue Xu, Mingchen Gao, Mario Buty, and Daniel J Mollura. Deep vessel tracking: A generalized probabilistic approach via deep learning. In 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pages 1363–1367. IEEE, 2016.
- [45] Jaeyoun Yi and Jong Beom Ra. A locally adaptive region growing algorithm for vascular segmentation. *International Journal of Imaging Systems and Technology*, 13(4):208–214, 2003.