Conformable Convolution for Topologically Constrained Learning of Complex Anatomical Structures

Yousef Yeganeh[1,2], **Göktuğ Güvercin**[1], **Nassir Navab**[1,2,3], **Azade Farshad**[4,5,1,2]
[1]Technical University of Munich, Germany
[2]Munich Center of Machine Learning, Germany
[3]ELLIS Unit Munich, Germany
[4]ELLIS Institute Finland
[5]Aalto University, Finland
y.yeganeh@tum.de

Abstract

While conventional computer vision emphasizes pixel-level and feature-based objectives, medical image analysis of intricate biological structures necessitates explicit representation of their complex topological properties. Despite their successes, deep learning models often struggle to accurately capture the connectivity and continuity of fine, sometimes pixel-thin, yet critical structures due to their reliance on implicit learning from data. To address this challenge, we introduce **Conformable Convolution**, a novel convolutional layer designed to explicitly impose topological consistency. Conformable Convolution learns adaptive kernel offsets that focus on regions of high topological significance within an image. This prioritization is guided by our proposed **Topological Posterior Generator** (TPG) module, which leverages persistent homology. The TPG module identifies key topological features and guides the convolutional layers by applying persistent homology to feature maps transformed into cubical complexes. Unlike existing approaches that are merely aware of topology, our method explicitly constrains the learning process to ensure topological correctness. The proposed modules are architecture-agnostic, enabling them to be integrated seamlessly into various architectures. We showcase the effectiveness of our framework in the segmentation task, where preserving the interconnectedness of structures is critical. The results on three diverse datasets demonstrate that our framework effectively preserves the topology both quantitatively and qualitatively.

Project Page — https://conformableconvolution.github.io

Introduction

Recent advances in medical image analysis, particularly in segmentation (He et al. 2023), have often prioritized pixel-level accuracy or visual quality, neglecting the inherent topological properties of anatomical structures. This oversight can lead to critical topological errors like false splits, merges, holes, or disconnected components, compromising the accuracy and reliability of analyses with potentially severe clinical consequences. For example, failing to accurately detect a ruptured vessel may lead to misdiagnosis of conditions such as aneurysms or stenoses. Therefore, ensuring realistic topological coherence is paramount in medical image analysis, where the continuity and connectivity of structures like vessels are essential. Although SOTA models (Hatamizadeh et al. 2021; He et al. 2023; Zhou, Yu, and Shi 2021) demonstrate strong performance in pixel-wise metrics, they often fail to capture these crucial topological characteristics. To address this gap, we introduce Conformable Convolution, an adaptive layer that explicitly constrains the feature learning using topological priors derived from persistent homology. Unlike existing topology-aware approaches that encourage correct topology through auxiliary losses or post-processing, our method embeds these constraints directly into the convolution operation itself, ensuring reliable preservation of anatomical connectivity throughout the learning process. The Conformable Convolution layers dynamically adjust sampling locations within their receptive field through learnable offsets, enabling the model to focus on regions of high topological interest. To identify these regions, we propose a novel Topological Posterior Generator (TPG) module that leverages persistent homology (Edelsbrunner, Letscher, and Zomorodian 2002) to quantify topological features across different scales – from connected components to loops and voids. By applying persistent homology to cubical complexes derived from feature maps, we obtain a discrete representation that effectively captures the underlying topology. Conformable Convolution layers are architecture-agnostic and seamlessly replace standard convolutions within existing architectures. This makes them easy to integrate into various models (as demonstrated in combination with four different types of architectures) to enforce topological preservation across diverse medical image analysis tasks, including segmentation.

We evaluate our framework on three diverse medical imaging datasets, where the continuity and connectivity of the structures are essential. Our framework effectively adheres to the topology in the input images, improving segmentation performance both qualitatively and quantitatively through conventional pixel-level segmentation metrics as well as connectivity-based metrics. The results of our evaluation on CHASE_DB1 for retinal vessel segmentation, HT29 for colon cancer cell segmentation, and ISBI12 for neuron electron microscopy segmentation demonstrate the ef-

fectiveness of the proposed modules in different shapes and structures. Furthermore, we propose a new evaluation metric through blood flow simulation to show the effectiveness of our model in vascular structures, which is presented in the supplementary materials. To summarize our main contributions: (1) We propose Conformable Convolution, which are convolutional layers with adaptively adjustable kernels guided by topological priors; (2) We propose the Topological Posterior Generator (TPG) module, which extracts the topological regions of interest for guiding the Conformable Convolution, (3) Our proposed modules are architecture-agnostic and can replace any convolution-based layer, (4) The quantitative and qualitative results of our experiments on the segmentation downstream task on three different organs and structures demonstrate the high impact of the proposed modules in topological metrics while achieving comparable or higher performance in pixel level metrics, (5) The source code of this work will be published upon its acceptance, and is available in the supplement.

Related Works

Previous work on topology-preserving methods can be broadly categorized into topology-aware networks and topology-aware objective functions (Wasserman 2018). Additionally, we cover methods that can potentially preserve topological structures and are relevant to our design.

Topology-preserving Layers and Networks Hofer et al. (2017) designed an input layer for a network that enables topological signatures as input and learning the optimal representations during training. Wyburd et al. (2021) utilizes the transformer-based VoxelMorph (Balakrishnan et al. 2019) learning to deform a topologically correct prior into the actual segmentation mask. SCOPE (Yeganeh et al. 2023) proposes a graph-based method that preserves continuity in retinal image segmentation. TA-Net (Wang, Xian, and Vakanski 2022) introduces a topology-aware network and utilizes medial axis transformation to encode the morphology of densely clustered gland cells in histopathological image segmentation. Gupta et al. (2022) employ a structure-guided approach to learn anatomical interactions, thereby facilitating the differentiation of tissues in medical segmentation. Horn et al. (2021) presents a topological layer in Graph Neural Networks. Gupta (Gupta et al. 2024) employs Discrete Morse Theory (DMT) (Forman 2002) for structural uncertainty estimation in Graph Convolution Networks (GCN) (Kipf and Welling 2016). Nishikawa, Ike, and Yamanishi (2024) applies Persistent Homology for point cloud analysis. Yi et al. (2024) proposes geometric-aware modeling for topology preservation in scalp electroencephalography (EEG). Moor et al. (2020) guides the bottleneck layer of an Autoencoder to produce topologically relevant features. Similar to their method, our method can be applied most effectively to the bottleneck to produce highly topological and faithful features.

Topology-preserving Objectives ToPoLoss (Hu et al. 2019; Clough et al. 2020) minimizes the Wasserstein distance in the *persistence diagram* (Vaserstein 1969; Cohen-Steiner et al. 2010) between the prediction and the ground truth. Stucki (Stucki et al. 2023) further improves Wasserstein matching by adopting the induced matching method on persistence barcodes. Prior to that, Centerline Dice (clDice) (Shit et al. 2021) was proposed as a tubular structure-based metric and loss function that improves segmentation results by incorporating accurate connectivity information. Another topology-aware objective function is the DMT loss (Hu et al. 2021), which helps detect saddle points that aid in reconstructing topologically incorrect regions. Hu (Hu 2022) computes warping errors at the homotopic level to promote topology. Recently, cbLoss (Shi et al. 2024) was introduced to mitigate data imbalance in medical image segmentation.

Adaptive and Structure-aware Layers Deformable convolution networks (DCN) (Dai et al. 2017) were first proposed to learn kernels that deform towards structures and shapes. Follow-up versions of DCN (Zhu et al. 2019; Wang et al. 2023; Xiong et al. 2024; Yeganeh, Farshad, and Navab 2023) expand on this idea by introducing additional deformations, incorporating it into foundation models, and further enhancing efficiency. With principles from DCN (Dong et al. 2022; Yang et al. 2022; Jin et al. 2019) adapt the shape and geometry of anatomical structures dynamically. Y-Net (Farshad et al. 2022) employed fast-Fourier convolutions to extract spectral features from medical images, inherently capturing underlying anatomical structures. Dynamic Snake Convolutions (DSC) (Qi et al. 2023) proposed snake-like kernels for deformable convolutions for topologically faithful tubular structure segmentation. However, the pre-set kernel shapes in DSC might neglect the performance while preserving the topology in other general shapes of structures. We, however, adopt a different strategy in topology preservation with an adaptive kernel; instead of pre-setting kernel shape, we aim to guide the kernel with offsets towards regions of higher topological interest.

Background

Topological Data Analysis (TDA) (Wasserman 2018) is a branch of applied mathematics focused on extracting meaningful geometric and topological features from high-dimensional, often noisy, and sparse data. Given a dataset $X \subset \mathbb{R}^n$, TDA focuses on analyzing the topological space (X, Θ), where Θ is an appropriate topology that captures the inherent structure of the data. Central to TDA is persistent homology, a technique that identifies and tracks topological features such as connected components, loops, and voids across multiple scales. These features are represented using simplicial complexes (K) or cubical complexes (Q), constructed from basic geometric shapes like points, lines, and triangles. These complexes serve as a bridge between the raw data (X) and its topological structure, which is quantified by homology groups such as Simplicial Complex: $K = \bigcup_{i=0}^{d} \sigma_i$, where σ_i are simplices or Cubical Complex: $Q = \bigcup_{i=0}^{d} c_i$, where c_i are cubes (Chazal and Michel 2021). TDA's capacity to derive robust, qualitative insights from complex data has led to its application in various fields, including biology, neuroscience, materials science, and social network analysis (Moor et al. 2020; Rieck et al. 2020).

In 2D medical imaging, **Cubical Complexes** are partic-

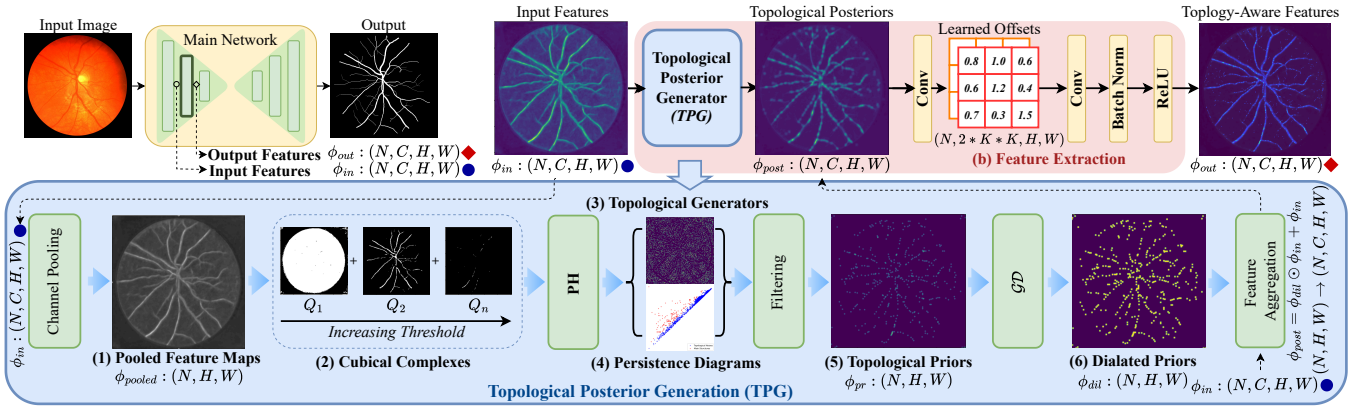

Figure 1: Conformable can replace other layers at different positions to extract topologically relevant features, as shown on the top-left. Our proposed layer receives the input feature map ϕ_{in} from the previous layer and generates ϕ_{post} by (a) *Topological Posterior Generation* component; then with the produced posteriors(ϕ_{post}) as the input, in (b) the first convolution layer generates offsets for the adaptive kernel of the second convolution.

ularly suitable due to the grid-like structure of the images (Santhirasekaram et al. 2023). Formally, a cubical complex Q in a 2D binary image consists of 0-Dimensional Cubes (0-Cells): Foreground pixels, denoted as $c_0 \in Q$, and 1-Dimensional Cubes (1-Cells): Connections between foreground pixels, denoted as $c_1 \in Q$. For our specific task, we focus on 0-dimensional cubes as the primary representation within the cubical complex. Persistent Homology **(PH)** tracks the evolution of these topological features (0-cells in our case) across a filtration of the cubical complex. Given a feature map ϕ and a threshold τ, the function $f_\tau(\phi) = Q$ maps ϕ to a cubical complex Q. Varying the threshold τ yields a nested sequence of cubical complexes:

$$\emptyset = Q_0 \in Q_1 \in Q_2 \in ... \in Q_n = Q \tag{1}$$

Persistence Diagram As **PH** is applied, one structure will be born (appear) and die (merge into other structures). Persistence Diagram (**PD**) documents the corresponding filtering threshold τ while a structure is born and dies. If a structure is born at τ_i and dies at τ_j, the tuple (τ_i, τ_j) would be recorded in PD. Here, we denote PD as the set containing all such tuples $\{(\tau_i, \tau_j)\}$ and a function $pers(.)$ to compute the *persistence* of a tuple (τ_i, τ_j):

$$pers(\tau_i, \tau_j) = |\tau_i - \tau_j| \tag{2}$$

Topological Generators In 2D images, topological generators are the pixel coordinates where significant topological events (birth or death of 0-cells) occur. They visually represent the starts and ends of distinct structures in an image. Figure 4-(b) showcases the positions of generators in orange pixels. Since PD documents the born-and-dead tuple of filtering threshold τ, we can define a function g that maps the set PD, which contains tuples of thresholds (τ_i, τ_j), to a set G, which contains nested tuples of pixel coordinates $((x_i, y_i), (x_j, y_j))$. So the set G contains all topological generators.

$$g : PD \mapsto G, \qquad g((\tau_i, \tau_j)) = ((x_i, y_i), (x_j, y_j)) \tag{3}$$

We provide a simplified visualization of the **PH** process in Figure 2, where a nested set of Q is generated using PH. The vessel has a longer lifespan since it spans a larger range of τ compared to noise, and according to Equation (2), the vessel has longer *persistence*. This demonstrates that noise generally has shorter persistence, allowing us to filter it in our methodology.

Methodology

This section presents our topology-guided conformable convolution layer and our proposed approach to utilize **PH**. Figure 1 shows the main components of our proposed Conformable layer, which can replace other layers, such as convolution, in a plug-and-play manner within a network.

Consider a semantic segmentation network θ, taking an input image $I \in \mathbb{R}^{B,C',H',W'}$, and producing a predicted segmentation map $y' = \theta(I)$. Given the ground truth segmentation map y, the network's objective is to minimize the Dice loss (Milletari, Navab, and Ahmadi 2016) between y and y'. Our topological module can process both raw images and intermediate feature maps; therefore, it can be inserted at any intermediate layer θ_i within the network θ. When inserted as the first layer (θ_0), the module operates directly on the input image I. For subsequent layers (θ_i ,$i > 0$), the module processes the feature map output of the preceding layer. For notational simplicity, we refer to the input to the module generically as a feature map.

Topological Posterior Generation

We are given an input feature map $\phi_{in} \in \mathbb{R}^{N \times C \times H \times W}$, where N, C, H, and W are the batch size, channels, height, and width, respectively. Our *TPG* block computes a weighted prior ϕ_{pr} that emphasizes regions with high topological interests, then aggregates the original semantics from ϕ_{in} back to the topological posterior ϕ_{post} which will be passed to the adaptive feature extractor (Figure 1-(b)).

In TPG, a channel pool layer denoted by ψ is applied to ϕ_{in} to extract the global patterns and to reduce the chan-

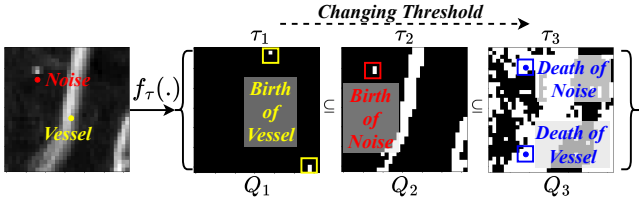

Figure 2: A visualization of how **PH** applies the filtering function $f_\tau(.)$ to an example image with vessel ■ and noise ■ by changing τ (τ_1, τ_2, τ_3), obtaining a nested set of cubical complexes Q (Q_1, Q_2, Q_3). As τ increases from τ_1 to τ_2, vessel ■ is first born at Q_1 and noise is later born at Q_2 ■. Both of them ■ die at Q_3, as τ further raises to τ_3.

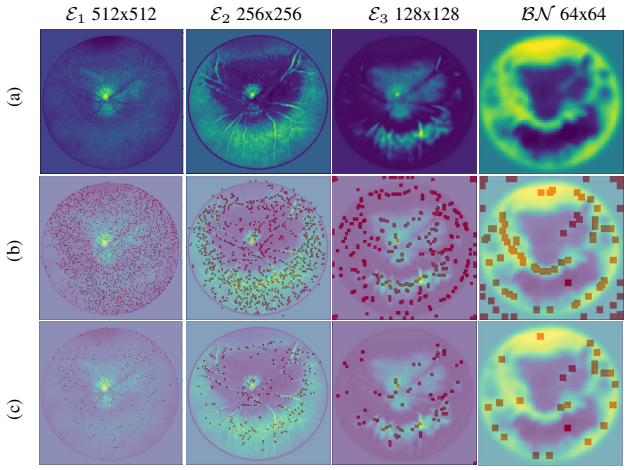

Figure 3: Topological Prior Visualization in *UNet + Conform* encoder ($\mathcal{E}$) and bottleneck ($\mathcal{BN}$) layers, showing (a) mean feature map, (b) all generators, and (c) feature priors after filtering.

nel dimensionality (*cf.* Figure 1(a-1)), getting $\phi_{pooled} \in \mathbb{R}^{N\times H\times W}$:

$$\phi_{pooled} = \psi(\phi_{in}) \tag{4}$$

As described in the background section, **PH** is later applied to ϕ_{pooled} to generate a set of tuples $\{(\tau_i, \tau_j) \mid (\tau_i, \tau_j) \in PD\}$, representing the birth and death times of topological features. Equation (3) then maps these tuples to a corresponding set of generators, denoted as G. Figure 1-(a-3) illustrates an example of G for a single ϕ_{pooled}, highlighting the presence of numerous redundant and noisy generators. As shown in our ablation study (Table 4), this unfiltered noise can negatively impact the topological faithfulness of the representation.

Filtering Generators As Edelsbrunner, Letscher, and Zomorodian (2002) suggests, structures with low persistence values often represent noise. To address this, we filter the set of generators G, retaining only those associated with significant topological features. We denote this filtered set as G_M. Formally, given a pair $(\tau_i, \tau_j) \in PD$ and a filtering

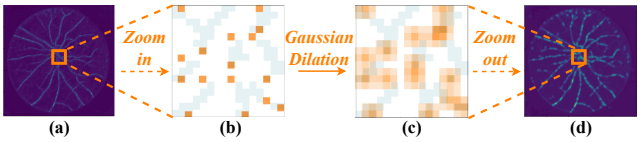

Figure 4: Gaussian dilation process in real and zoomed-in feature map: (a) ϕ_{pr} in a vessel feature map; (b) a zoomed-in synthetic feature map, depicting ϕ_{pr} emphasizing on regions of high topological interests, (c) the effect of Gaussian dilation in dilating the topologically significant regions; (d) the impact of Gaussian dilation on the vessel feature map.

threshold τ_0, we compute:

$$\mathbb{I}(\tau_i, \tau_j) = \begin{cases} 1 & \text{if } pers(\tau_i, \tau_j) > \tau_0, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The indicator function $\mathbb{I}(.)$ allows us to construct a binary mask M over the entire *PD*. Through element-wise multiplication, we obtain the filtered generators G_M:

$$M = \{\mathbb{I}(\tau_i, \tau_j) \mid (\tau_i, \tau_j) \in PD\}, \qquad G_M = M \odot G \tag{6}$$

Generating Topological Priors Since G_M contains a set of coordinates of generators that mark the start and ending points of any connected components, regions with concentrated generators should be of high topological interest. The next step will be to convert these coordinates into a weighted prior, encoding the topological information into the learned offsets field, which will later be utilized for feature extraction. Such conversion from G_M to ϕ_{pr} can be easily achieved by first constructing a zero $\phi_{pr} \in \mathbb{R}^{B\times H\times W}$, then filling the (i, j) entry with ones if such entry is in G_M:

$$\phi_{pr}(i,j) = \begin{cases} 1 & \forall (i,j) \in G_M \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

A visualization of the topological prior at different network layers is provided in Figure 3.

Gaussian Dilation The obtained binary ϕ_{pr} is indeed weighting regions with high topological interests. As depicted in Figure 4-(b), ϕ_{pr} could effectively capture the starting and ending points of a vessel and assign a weight to it. However, its pixel-wise nature makes it hard to cover all the disconnected regions. Therefore, we propose a Gaussian dilation strategy that formulates ϕ_{pr} as a probabilistic weighted prior. This is achieved by applying convolution to ϕ_{pr} with a 3×3 normalized Gaussian kernel denoted as $\mathcal{GD}$. We use $*$ to denote the convolution operator.

$$\phi_{\text{dil}} = \mathcal{GD} * \phi_{pr} \tag{8}$$

As shown in Figure 4-(c), we assign Gaussian distributions to all disconnected regions, which are of high topological interest. To visualize its effect in a real feature map, Figure 4-(a) and Figure 4-(d) show the effect after Gaussian dilation is applied. In Table 4, Gaussian dilation is also justified to contribute to the topological results.

Topological Posterior Generation ϕ_{pr} effectively augments topology significant parts. However, to prevent loss of

valuable information in topological sampling, as it is shown in Figure 1-(a), the dilated prior ϕ_{dil} is first used to augment the topological significant parts of original input ϕ_{in}, then it is aggregated with the ϕ_{in}, forming a stronger topological posterior estimation ϕ_{post}:

$$\phi_{post} = \phi_{dil} \odot \phi_{in} + \phi_{in} \tag{9}$$

Conformable Convolution

Inspired by layers with an adaptive kernel design, we propose *Conformable Convolutions* (Conform for short). Unlike standard convolution, convolutions with an adaptive kernel reposition convolutional kernels w_i using learnable offsets Δp_c. This adaptability allows the model to better focus on contours and interconnected segments through offset convolution $g(.)$. In standard convolution, a fixed grid R defines the receptive field and dilation of a kernel. The kernel elements, indexed by grid coordinates, are multiplied with corresponding pixel values from the input feature map $\phi_{in}(.)$. These products are then aggregated to produce each pixel p in the output feature map $\phi_{out}(.)$, as formulated:

$$\begin{aligned} R &= \{(-1,-1),(-1,0),...,(1,1)\}, \\ \phi_{out}(p) &= \sum_{p_c \in R} w_c \cdot \phi_{in}(p + p_c) \end{aligned} \tag{10}$$

Learnable offsets in convolution enable the kernel to sample pixel values from non-regular grid locations within the input feature map. This modulation is achieved through a set of offsets $\{\Delta p_c\}_{c=1}^{C}$, where $C = |R|$ represents the cardinality of the regular grid R on which the kernel operates.

$$\begin{aligned} \{\Delta p_c\}_{c=1}^{C} &= g(\phi_{in}), \\ \phi_{out}(p) &= \sum_{p_c \in R} w_c \cdot \phi_{in}(p + p_c + \Delta p_c) \end{aligned} \tag{11}$$

The modulation of these kernels is susceptible to artifacts and high contrast inside the receptive field. In topological posterior maps, those artifacts and contrasts are suppressed by using generated birth and death points with filtration mechanisms. In this way, the adjustable convolution is still applied to the input feature maps; nonetheless, the offset adjustment is refined by topological activity regions, which introduce a new offset space with topological deformation:

$$\begin{aligned} \{\Delta \hat{p}_c\}_{c=1}^{C} &= g(TPG(\phi_{in})), \\ \phi_{out}(p) &= \sum_{p_c \in R} w_c \cdot \phi_{in}(p + p_c + \Delta \hat{p}_c) \end{aligned} \tag{12}$$

Experiments and Results

In this section, we provide a comprehensive evaluation of our proposed layer for topology-aware segmentation on three distinct medical imaging datasets: CHASE, HT29, and ISBI12. First, we report the experimental setup. Then, we investigate the integration of our layer in various backbones and compare it with other state-of-the-art layers designed specifically for modeling geometry and topology. Then, we follow a similar strategy, yet compare it against different baselines. Finally, we present an ablation study of the components in our layer configuration. The implementation details are reported in the supplement. In all experiments, we report the mean and standard deviation across multiple runs with different random seeds, unless otherwise specified.

Experimental Setup

Datasets We evaluate our model on three datasets with diverse topological properties. The *ISBI12* dataset (Arganda-Carreras et al. 2015) featuring intricate network-like structures of neurons with numerous loops and connections, presents a significant challenge for preserving both *0-dim* topology (# of disconnected components) as well as *1-dim* topology (# of holes). In contrast, *CHASE_DB1* retinal vessel dataset (Fraz et al. 2012), consisting of 28 images, lacks loops but exhibits complex vessel structures that demand accurate preservation of connected components (*0-dim* topology). The *HT29* colon cancer cell dataset from the Broad BioImage Benchmark Collection (Carpenter et al. 2006; Ljosa, Sokolnicki, and Carpenter 2012), also known as *BBBC*, is characterized by blob-like foreground structures with few holes, making it less sensitive to errors in *1-dim* topological error, such as the $error_{\beta_1}$.

Evaluation Metrics Standard classification metrics assess individual pixels within segmented regions without considering their structural relationships or connectivity. To investigate the topological properties of segmentation maps across different homology groups, as a central goal of this paper, we employ four topological and two entropy-based evaluation metrics. Specifically, we utilize *clDice* (Shit et al. 2021) to evaluate the center-line continuity of tubular structures. We use *Betti-0* (β_0) and *Betti-1* (β_1) (Vietoris 1927) to count the number of connected components and independent holes, respectively. The *Euler characteristic* (χ) serves as a topological invariant metric, quantifying the shape of the segmentation manifold that encompasses all possible topological spaces of the segmented regions. We employ the *Adjusted Rand Index (ARI)* (Arganda-Carreras et al. 2015) to measure the similarity of randomly chosen pixel pairs belonging to the same or different segmented regions, and the *Variation of Information (VI)* (Meilă 2007) to quantify the amount of information that a cluster contains about the other one. In addition to these topology-focused metrics, we report the commonly utilized pixel-wise segmentation metrics: the area under the curve (*AUC*) and *Dice Score* between the ground truth and predicted segmentation maps.

Results

We validate the performance gain achieved by incorporating our proposed layer with simple baselines compared to SOTA segmentation models in Table 1 on the CHASE dataset, Table 3 on ISBI12, and qualitatively in Figure 3 in the supplement. In pixel-wise metrics, SGL (Zhou, Yu, and Shi 2021) and FR-UNet (Liu et al. 2022) achieve the most promising results; nevertheless, they have difficulty perceiving inter-pixel connections and the topology of segmented vessel branches. In terms of continuity and topology preservation, SCOPE (Yeganeh et al. 2023) and the Conformable layer

Architecture	Pixel-Wise Metrics		Topological Metrics					
	AUC (%) ↑	Dice (%) ↑	clDice ↑	err_{β_0} ↓	err_{β_1} ↓	err_{χ} ↓	ARI ↓	VI ↓
SOTA General Segmentation Models								
SwinUNETR	92.2	75.8	75	37.4	3.5	38.1	0.20	0.36
SwinUNETR-V2	90.3	74.4	73	39.9	1.7	40.5	0.22	0.37
FR-UNet	<u>99.1</u>	<u>81.5</u>	73	61.0	2.8	64.4	—	—
SGL	**99.2**	**82.7**	75	42.6	2.3	46.0	—	—
+ Conform (Ours)	98.3	80.8	79	33.4	2.0	30.8	0.18	<u>0.29</u>
SOTA Topological Segmentation Models								
VGN	-	73.0	78	71.9	4.4	69.5	—	—
SCOPE + Dice	95.4	80.0	<u>80</u>	32.6	2.0	28.5	0.17	**0.28**
+ Conform (Ours)	96.6	79.2	**81**	29.5	**1.5**	24.9	<u>0.15</u>	0.30
SCOPE + clDice	98.8	80.2	**81**	24.2	<u>1.6</u>	<u>22.7</u>	**0.14**	0.30
+ Conform (Ours)	98.6	79.4	**81**	<u>21.5</u>	2.1	**19.8**	**0.14**	0.30
Baseline Segmentation Models w. and w/o Conform								
UNet	92.3	79.3	79	26.9	2.7	28.5	0.19	0.30
+ Conform (Ours)	94.2	79.7	**81**	21.6	2.1	**20.6**	<u>0.17</u>	**0.28**
Y-Net	98.0	78.0	76	27.9	3.1	24.4	0.18	0.31
+ Conform (Ours)	98.7	80.2	<u>79</u>	**21.1**	2.0	23.5	<u>0.17</u>	**0.28**

Table 1: Segmentation Performance Compared to SOTA Segmentation Models on CHASE (Fraz et al. 2012). The best and second-best performing methods are shown in **bold** and <u>underlined</u>, respectively.

Architecture	Pixel-Wise Metrics			Topological Metrics				
	AUC (%) ↑	Dice (%) ↑	clDice (%) ↑	err_{β_0} ↓	err_{β_1} ↓	err_{χ} ↓	ARI ↓	VI ↓
SwinUNETR	<u>92.7</u>±0.6	**81.1**±0.2	92.7 ±0.1	11.5 ±4.1	8.2 ±2.2	10.1 ±3.6	0.16 ±0.07	0.81 ±0.02
SwinUNETR-V2	**92.9** ±0.4	<u>80.9</u> ±0.3	92.9 ±0.2	10.3 ±3.1	**7.5** ±2.5	9.6 ±3.9	0.15 ±0.04	0.80 ±0.02
SCOPE	-	79.8±0.0	**93.9**±0.0	**6.2**±0.0	9.3±0.0	9.5±0.0	-	-
UNet	91.3±0.8	78.7±1.6	92.1±1.7	14.4±3.8	23.5±4.1	24.8±3.7	0.17±0.04	0.85±0.11
+ Conform	92.4±1.5	80.6±0.9	**93.9**±0.6	13.0±3.7	<u>7.9</u>±2.9	**8.4**±2.9	**0.15**±0.00	**0.79**±0.00
Y-Net	91.1±1.7	78.3±2.5	92.5±1.4	17.0±4.4	14.8±4.6	14.0±5.8	0.18±0.03	0.85±0.10
+ Conform	91.6±1.2	79.6±1.2	93.6±0.6	10.5±3.9	<u>7.9</u>±2.7	8.4±3.7	0.15±0.02	0.81±0.07

Table 2: Segmentation Performance Compared to SOTA Segmentation Models on ISBI12 (Arganda-Carreras et al. 2015). The best and second best performing methods are shown as **bold** and <u>underlined</u>, respectively.

with Y-Net achieve the best results, as validated by our qualitative results shown in Figure 3 of the supplement. The Conformable layer leverages topological awareness in Y-Net, providing a noticeable contribution to topological segmentation compared to its standard version. However, possibly due to the size of the model, there is no observable improvement in UNet (Ronneberger, Fischer, and Brox 2015). VGN (Shin et al. 2019) is, on the contrary, liable to do over-segmentation in which curvilinear structures can be topologically segmented, yet additional isolated vessel islands would also be generated. This leads to many disconnected regions in the prediction map, thereby decreasing the dice and connectivity scores. It should be noted that although SCOPE achieves higher performance in some topological metrics, its architecture is designed to tackle the task at hand. Our Conformable layer, on the other hand, is architecture-agnostic and can be combined with different models.

We further compare the effectiveness of the Conformable layer to other adaptable kernel layers on three anatomically dense medical imaging datasets, spanning microscopic anatomical structures to larger yet more entangled ones. As shown in Table 3, while comparing with deformable convolution (Dai et al. 2017) and the SOTA Dynamic Snake Convolution (DSC) (Qi et al. 2023), we observe that although Conformable yields modest trade-offs in traditional segmentation metrics (AUC and Dice scores), it excels significantly in continuity-preserving metrics. Notably, Conformable achieves substantially lower error rates in topological measurements—reducing error_{χ} from 13.3 to 6.1 (a 54.1% decrease compared to Deform) and from 20.58 to 6.1 (a 70.4% decrease compared to DSC) on the HT29 dataset. On the ISBI12 dataset, Conformable outperforms both benchmarks across most metrics, demonstrating improved segmentation quality ($+1.3\%$ Dice vs. DSC) while maintaining better structural continuity (with error_{χ} reduced from 12.6 to 8.4, a 33.3% improvement).

Layer	Pixel-Wise Metrics		Topological Metrics					
	AUC (%) ↑	Dice (%) ↑	clDice (%) ↑	err_{β_0} ↓	err_{β_1} ↓	err_{χ} ↓	ARI ↓	VI ↓
HT29 (Carpenter et al. 2006; Ljosa, Sokolnicki, and Carpenter 2012)								
Deform	**99.6**±0.2	95.8±2.1	**93.7**±4.0	8.2±3.6	13.1±4.7	13.3±4.2	0.05 ±0.03	**0.19** ±0.02
DSC	99.4 ±0.3	**95.8** ±2.0	87.6 ±3.4	8.9 ±2.8	**7.8** ±3.1	20.6 ±7.2	0.06 ±0.07	0.21 ±0.01
Conform (Ours)	99.1 ±0.6	94.6 ±1.3	93.1 ±4.5	**5.9** ±2.4	9.6 ±3.1	**6.1** ±2.2	**0.04** ±0.01	**0.19** ±0.06
ISBI12 (Arganda-Carreras et al. 2015)								
Deform	91.4 ±0.9	79.4 ±1.4	93.3 ±0.8	15.5 ±3.6	8.9 ±3.0	13.6 ±5.0	0.16 ±0.1	0.82 ±0.0
DSC	91.6 ±0.2	79.6 ±1.5	93.2 ±0.1	13.2 ±4.5	9.7 ±7.0	12.6 ±2.8	0.17 ±0.0	0.82 ±0.0
Conform (Ours)	**92.4** ±1.5	**80.6** ±0.9	**93.9** ±0.6	**13.0** ±3.7	**7.9** ±2.9	**8.4** ±2.9	**0.15** ±0.0	**0.79** ±0.0
CHASE (Fraz et al. 2012)								
Deform	94.0 ±0.3	79.3 ±0.1	78.6 ±0.3	24.1 ±1.7	2.79 ±0.2	25.5 ±2.8	0.18 ±0.00	**0.28** ±0.00
DSC	**95.9** ±0.2	79.6 ±0.2	79.9 ±0.4	28.3 ±1.7	3.67 ±0.5	26.4 ±1.4	0.18 ±0.00	0.30 ±0.00
Conform (Ours)	94.2 ±0.2	**79.7** ±0.4	**80.6** ±0.0	**21.6** ±3.0	**2.20** ±0.4	**20.9** ±3.6	**0.17** ±0.00	**0.28** ±0.00

Table 3: Segmentation Performance Compared to SOTA Layers. DSC (Qi et al. 2023) and Deform (Dai et al. 2017) on CHASE, HT29, and ISBI12. The layers are inserted at the bottleneck of a UNet model.

Fil.	$\mathcal{GD}$	Aggr.	clDice ↑	err_{β_0} ↓	err_{β_1} ↓	err_{χ} ↓	ARI ↓	VI ↓
-	✓	✓	79±0.00	32.7±1.1	3.2±0.5	33.8±1.5	0.19±0.00	0.30±0.02
✓	-	✓	79±0.01	23.4±1.4	3.0±0.3	23.8±2.1	0.19±0.01	0.28±0.01
✓	✓	-	80±0.00	24.8±0.9	2.9±0.6	25.2±1.3	0.18±0.03	0.29±0.04
✓	✓	✓	**81**±0.00	**21.6**±3.0	**2.1**±0.4	**20.6**±3.6	**0.17**±0.00	**0.28**±0.00

Table 4: Ablation Study of Different Components on CHASE (Fraz et al. 2012). The model with all components corresponds to "UNet + Conform" in Table 1. $\mathcal{GD}$: Gaussian Dilation, **Fil.**: Filtration, **Aggr.**: Feature Aggregation

Effect of Different Components To further justify our design choice of methodology, we ablate the filtration, Gaussian dilation, and feature aggregation process to learn their effects on the results. As shown in Table 4, when no filtering is applied to the generators in TPG, regions with noise would not be filtered and would be assigned a high weight in f_{prior}. This also leads to noise in the final prediction, causing worse topological metrics. When we remove the Gaussian dilation module in Table 4, the topological results also worsen. This indicates that Gaussian dilation can enhance the local features with topological significance, thereby improving the final segmentation results. At last, we also block the aggregation of input feature maps to see if the fusion of semantics from ϕ_{in} is really effective. After the aggregation is blocked, the equation is updated to:

$$\phi_{post} = \phi_{dil} \odot \phi_{in} \tag{13}$$

We show that such an aggregation from ϕ_{in} could benefit the gradient flow and is also beneficial for the topological segmentation results in Table 4. Overall, our proposed design produces segmentations with fewer clustering errors and accurate topological consistency, striking a balance between precise boundary delineation and the preservation of critical structural properties—making it particularly valuable for applications where maintaining the connectivity of thin, elongated structures is crucial. The Conformable layer learns a better representation of geometry and anatomical consistency for continuity preservation, while not sacrificing pixel-wise accuracy. This significantly enhances Betti and Euler metrics, contributing to the similarity of cluster segments (ARI and VI) and center-line connectivity (clDice), due to the amplified wholeness of anatomical structures. In addition, we ablate the position of inserting and the number of Conformable layers in the architecture in the appendix.

Conclusion

In this work, we introduce Conformable Convolution, a novel, architecture-agnostic layer that bridges topological principles and neural networks. By explicitly encoding topological constraints into the learning process, without compromising pixel-level performance, the critical structural properties are preserved. Our approach incorporates a topological posterior generator (TPG) module, which identifies and prioritizes regions of high topological significance within feature maps. By integrating persistent homology, we ensure the preservation of critical topological features, such as connectivity and continuity, which are often overlooked by conventional deep learning models. Our proposed modules are designed to be architecture-agnostic, allowing seamless integration into various existing networks. Through extensive experiments on diverse medical imaging datasets, we demonstrate the effectiveness of our framework in adhering to the topology and improving segmentation performance, both quantitatively and qualitatively.

Acknowledgments

This work is partially supported by the Munich Center for Machine Learning in collaboration between the ELLIS Munich Unit and the ELLIS Institute Finland.

References

Arganda-Carreras, I.; Turaga, S. C.; Berger, D. R.; Cireşan, D.; Giusti, A.; Gambardella, L. M.; Schmidhuber, J.; Laptev, D.; Dwivedi, S.; Buhmann, J. M.; et al. 2015. Crowdsourcing the creation of image segmentation algorithms for connectomics. *Frontiers in neuroanatomy*, 9: 142.

Balakrishnan, G.; Zhao, A.; Sabuncu, M. R.; Guttag, J.; and Dalca, A. V. 2019. VoxelMorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 38(8): 1788–1800.

Carpenter, A. E.; Jones, T. R.; Lamprecht, M. R.; Clarke, C.; Kang, I. H.; Friman, O.; Guertin, D. A.; Chang, J. H.; Lindquist, R. A.; Moffat, J.; et al. 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome biology*, 7: 1–11.

Chazal, F.; and Michel, B. 2021. An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Frontiers in artificial intelligence*, 4: 108.

Clough, J. R.; Byrne, N.; Oksuz, I.; Zimmer, V. A.; Schnabel, J. A.; and King, A. P. 2020. 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.

Cohen-Steiner, D.; Edelsbrunner, H.; Harer, J.; and Mileyko, Y. 2010. Lipschitz functions have L p-stable persistence. *Foundations of computational mathematics*, 10(2): 127–139.

Dai, J.; Qi, H.; Xiong, Y.; Li, Y.; Zhang, G.; Hu, H.; and Wei, Y. 2017. Deformable convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, 764–773.

Dong, S.; Pan, Z.; Fu, Y.; Yang, Q.; Gao, Y.; Yu, T.; Shi, Y.; and Zhuo, C. 2022. DeU-Net 2.0: Enhanced deformable U-Net for 3D cardiac cine MRI segmentation. *Medical Image Analysis*, 78: 102389.

Edelsbrunner; Letscher; and Zomorodian. 2002. Topological persistence and simplification. *Discrete & Computational Geometry*, 28: 511–533.

Farshad, A.; Yeganeh, Y.; Gehlbach, P.; and Navab, N. 2022. Y-Net: A spatiospectral dual-encoder network for medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 582–592. Springer.

Forman, R. 2002. A user's guide to discrete Morse theory. *Séminaire Lotharingien de Combinatoire [electronic only]*, 48: B48c–35.

Fraz, M. M.; Remagnino, P.; Hoppe, A.; Uyyanonvara, B.; Rudnicka, A. R.; Owen, C. G.; and Barman, S. A. 2012. An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Transactions on Biomedical Engineering*, 59(9): 2538–2548.

Gupta, S.; Hu, X.; Kaan, J.; Jin, M.; Mpoy, M.; Chung, K.; Singh, G.; Saltz, M.; Kurc, T.; Saltz, J.; et al. 2022. Learning topological interactions for multi-class medical image segmentation. In *European Conference on Computer Vision*, 701–718. Springer.

Gupta, S.; Zhang, Y.; Hu, X.; Prasanna, P.; and Chen, C. 2024. Topology-aware uncertainty for image segmentation. *Advances in Neural Information Processing Systems*, 36.

Hatamizadeh, A.; Nath, V.; Tang, Y.; Yang, D.; Roth, H. R.; and Xu, D. 2021. Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In *International MICCAI Brainlesion Workshop*, 272–284. Springer.

He, Y.; Nath, V.; Yang, D.; Tang, Y.; Myronenko, A.; and Xu, D. 2023. Swinunetr-v2: Stronger swin transformers with stagewise convolutions for 3d medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 416–426. Springer.

Hofer, C.; Kwitt, R.; Niethammer, M.; and Uhl, A. 2017. Deep learning with topological signatures. *Advances in neural information processing systems*, 30.

Horn, M.; De Brouwer, E.; Moor, M.; Moreau, Y.; Rieck, B.; and Borgwardt, K. 2021. Topological graph neural networks. *arXiv preprint arXiv:2102.07835*.

Hu, X. 2022. Structure-aware image segmentation with homotopy warping. *Advances in Neural Information Processing Systems*, 35: 24046–24059.

Hu, X.; Li, F.; Samaras, D.; and Chen, C. 2019. Topology-preserving deep image segmentation. *Advances in neural information processing systems*, 32.

Hu, X.; Wang, Y.; Fuxin, L.; Samaras, D.; and Chen, C. 2021. Topology-aware segmentation using discrete Morse theory. *arXiv preprint arXiv:2103.09992*.

Jin, Q.; Meng, Z.; Pham, T. D.; Chen, Q.; Wei, L.; and Su, R. 2019. DUNet: A deformable network for retinal vessel segmentation. *Knowledge-Based Systems*, 178: 149–162.

Kipf, T. N.; and Welling, M. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

Liu, W.; Yang, H.; Tian, T.; Cao, Z.; Pan, X.; Xu, W.; Jin, Y.; and Gao, F. 2022. Full-resolution network and dual-threshold iteration for retinal vessel and coronary angiograph segmentation. *IEEE Journal of Biomedical and Health Informatics*, 26(9): 4623–4634.

Ljosa, V.; Sokolnicki, K. L.; and Carpenter, A. E. 2012. Annotated high-throughput microscopy image sets for validation. *Nature methods*, 9(7): 637–637.

Meilă, M. 2007. Comparing clusterings—an information based distance. *Journal of multivariate analysis*, 98(5): 873–895.

Milletari, F.; Navab, N.; and Ahmadi, S.-A. 2016. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, 565–571. Ieee.

Moor, M.; Horn, M.; Rieck, B.; and Borgwardt, K. 2020. Topological autoencoders. In *International conference on machine learning*, 7045–7054. PMLR.

Nishikawa, N.; Ike, Y.; and Yamanishi, K. 2024. Adaptive Topological Feature via Persistent Homology: Filtration Learning for Point Clouds. *Advances in Neural Information Processing Systems*, 36.

Qi, Y.; He, Y.; Qi, X.; Zhang, Y.; and Yang, G. 2023. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 6070–6079.

Rieck, B.; Yates, T.; Bock, C.; Borgwardt, K.; Wolf, G.; Turk-Browne, N.; and Krishnaswamy, S. 2020. Uncovering the topology of time-varying fMRI data using cubical persistence. *Advances in neural information processing systems*, 33: 6900–6912.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, 234–241. Springer.

Santhirasekaram, A.; Winkler, M.; Rockall, A.; and Glocker, B. 2023. Topology Preserving Compositionality for Robust Medical Image Segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 543–552.

Shi, P.; Hu, J.; Yang, Y.; Gao, Z.; Liu, W.; and Ma, T. 2024. Centerline Boundary Dice Loss for Vascular Segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 46–56. Springer.

Shin, S. Y.; Lee, S.; Yun, I. D.; and Lee, K. M. 2019. Deep vessel segmentation by learning graphical connectivity. *Medical image analysis*, 58: 101556.

Shit, S.; Paetzold, J. C.; Sekuboyina, A.; Ezhov, I.; Unger, A.; Zhylka, A.; Pluim, J. P.; Bauer, U.; and Menze, B. H. 2021. clDice-a novel topology-preserving loss function for tubular structure segmentation. In *Proceedings of CVPR*, 16560–16569.

Stucki, N.; Paetzold, J. C.; Shit, S.; Menze, B.; and Bauer, U. 2023. Topologically faithful image segmentation via induced matching of persistence barcodes. In *International Conference on Machine Learning*, 32698–32727. PMLR.

Vaserstein, L. N. 1969. Markov processes over denumerable products of spaces, describing large systems of automata. *Problemy Peredachi Informatsii*, 5(3): 64–72.

Vietoris, L. 1927. Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen. *Mathematische Annalen*, 97(1): 454–472.

Wang, H.; Xian, M.; and Vakanski, A. 2022. Ta-net: Topology-aware network for gland segmentation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 1556–1564.

Wang, W.; Dai, J.; Chen, Z.; Huang, Z.; Li, Z.; Zhu, X.; Hu, X.; Lu, T.; Lu, L.; Li, H.; et al. 2023. Internimage: Exploring large-scale vision foundation models with deformable convolutions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14408–14419.

Wasserman, L. 2018. Topological data analysis. *Annual Review of Statistics and Its Application*, 5: 501–532.

Wyburd, M. K.; Dinsdale, N. K.; Namburete, A. I.; and Jenkinson, M. 2021. TEDS-Net: enforcing diffeomorphisms in spatial transformers to guarantee topology preservation in segmentations. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 250–260. Springer.

Xiong, Y.; Li, Z.; Chen, Y.; Wang, F.; Zhu, X.; Luo, J.; Wang, W.; Lu, T.; Li, H.; Qiao, Y.; et al. 2024. Efficient Deformable ConvNets: Rethinking Dynamic and Sparse Operator for Vision Applications. *arXiv preprint arXiv:2401.06197*.

Yang, X.; Li, Z.; Guo, Y.; and Zhou, D. 2022. DCU-net: A deformable convolutional neural network based on cascade U-net for retinal vessel segmentation. *Multimedia Tools and Applications*, 81(11): 15593–15607.

Yeganeh, Y.; Farshad, A.; and Navab, N. 2023. Anatomy-Aware Masking for Inpainting in Medical Imaging. In *International Workshop on Shape in Medical Imaging*, 35–46. Springer.

Yeganeh, Y.; Güvercin, G.; Xiao, R.; Abuzer, A.; Adeli, E.; Farshad, A.; and Navab, N. 2023. SCOPE: Structural Continuity Preservation for Retinal Vessel Segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 3–13. Springer.

Yi, K.; Wang, Y.; Ren, K.; and Li, D. 2024. Learning Topology-Agnostic EEG Representations with Geometry-Aware Modeling. *Advances in Neural Information Processing Systems*, 36.

Zhou, Y.; Yu, H.; and Shi, H. 2021. Study group learning: Improving retinal vessel segmentation trained with noisy labels. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24*, 57–67. Springer.

Zhu, X.; Hu, H.; Lin, S.; and Dai, J. 2019. Deformable convnets v2: More deformable, better results. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 9308–9316.