EXPTOPO: EXPLICIT TOPOLOGICAL MODELING FOR BRONCHUS SEGMENTATION

Mingyue Zhao[1,2], *Xiaolan Qiu*[3], *S.Kevin Zhou*[1,2,4,5]

[1] School of Biomedical Engineering, Division of Life Sciences and Medicine,
University of Science and Technology of China (USTC), China
[2] Center for Medical Imaging, Robotics, Analytic Computing & Learning (MIRACLE),
Suzhou Institute for Advance Research, USTC, China
[3]Suzhou Key Laboratory of Microwave, Imaging, Processing and Application Technology,
Suzhou Aerospace Information Research Institute, China
[4]Key Laboratory of Precision and Intelligent Chemistry,
University of Science and Technology of China (USTC), China
[5]Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences (CAS),
Institute of Computing Technology, China

ABSTRACT

Airway extraction is paramount in the early diagnosis and treatment of respiratory diseases. As a tree-like structure, both topological-aware learning and voxel-wise classification are equally crucial for the airway. However, existing methods demonstrate insufficient topological learning, emphasizing only the supervision of individual key topological points. Consequently, this paper proposes a **Exp**licit **Topo**logical Modeling (ExpTopo) approach to aid airway segmentation. It explicitly introduces topological metric space learning based on semantic segmentation, enhancing the model's structural perception by implementing global skeleton-level sparse topological learning (STL) and local voxel-level dense topological perception (DTP). Extensive experimental results demonstrate that the algorithm achieves competitive performance at both the topological and voxel levels. Code will be available in https://github.com/MorineZ/ExpTopo.

***Index Terms*—** Airway extraction, explicit topological modeling, sparse topological learning, dense topological perception

1. INTRODUCTION

Bronchial extraction is a paramount prerequisite for the early diagnosis and treatment of pulmonary diseases. It is the basis of virtual surgical navigation and bronchoscopic guidance. Morphological changes in the airways, particularly the remodeling of small bronchi, are often key structural indicators in the early stages of lung diseases [1].

Recently, deep learning methods [2–11] have become the mainstream approach for airway segmentation. A line of studies [8–10, 12] have focused on addressing class imbalance in segmentation tasks by optimizing network architectures or enhancing shallow gradient propagation. Nevertheless, this pure voxel-wise classification mode is still constrained by the limited receptive field and potentially causes branch breakage or the totally missed detection [13]. *1) Unlike convex structures, dense prediction of the tree-like airway structure requires both voxel-level classification and topology-aware learning to be equally emphasized.* Thus, some researchers [14–19] have attempted to explicitly incorporate topology-aware learning. They emphasize the continuity of segmentation by introducing an additional centerline prediction task [15, 18, 19] or an auxiliary skeleton-level loss [14, 16, 17]. However, *2) the existing methods demonstrate insufficient topological learning, emphasizing only the supervision of individual key topological points.* To this end, this paper proposes a novel **Exp**licit **Topo**logical Modeling (ExpTopo) approach, which introduces a topological representation space on top of the voxel space in a multi-head manner. In this topological space, following 1), we incorporate sparse topological learning to expand the receptive field at the topological level, while bridging the gap between voxel-level predictions and topology-level awareness. Regarding issue 2), we posit that **within difficult regions, such as branch discontinuities or occlusions, each voxel matters for dense voxel-level predictions and topological preservation.** As a result, we innovatively propose the concept of dense topological representation that robustly enhances local structural perception, thereby strengthening dense prediction.

Overall, the algorithm's main contributions are threefold: 1) In this study, skeleton-level sparse topology learning is employed to strengthen long-distance topological perception, addressing the constraints of limited receptive fields and improving the completeness of branch detection. 2) Inspired by

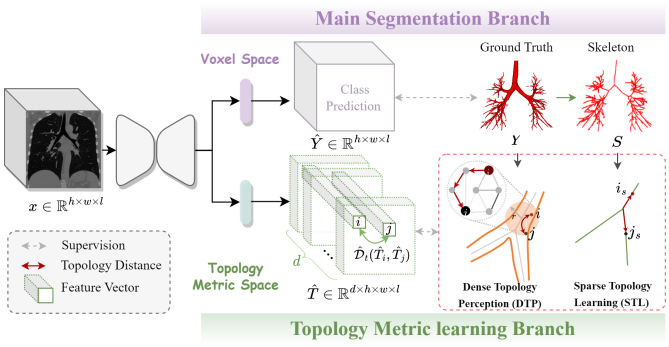

Fig. 1. Flowchart of our framework - ExpTopo.

the critical role of each voxel in hard regions, this study implements topology-aware dense pixel modeling by introducing dense topology metric to strengthen local structure perception and dense prediction. 3) The algorithm demonstrated superior performance across both datasets, particularly excelling in topological metrics, thereby validating its effectiveness.

2. METHODOLOGY

Fig. 1 illustrates the overall framework of the algorithm, which consists of a main segmentation branch and an auxiliary topology metric learning branch in a multi-head manner. The segmentation branch is dedicated to learning precise voxel-level airway prediction $\hat{Y}$ within the image space. The auxiliary branch is responsible for learning a reliable topological metric space $\hat{T} \in \mathbb{R}^{d\times h\times w\times l}$, where the distance between two points is proportional to the topological distance along points. The topological metric learning encompasses skeleton-level sparse topological learning (STL) and voxel-level dense topological perception (DTP).

2.1. Sparse Topological Learning (STL)

As depicted in Fig. 2I, given the skeleton of the airway tree S, we first perform branch pattern analysis to determine the parent-child relationships between branches. Based on this relationship graph, we calculate the distance between any two skeletal points i and j $(i \neq j, and\ i, j \in \Omega_{skel})$ along the skeleton, where the Ω_{skel} is the set of skeleton points and the distance is termed as *Sparse Topological Distance* $\mathcal{D}_{st}$. When i and j are located on the same branch or adjacent parent-child branches, they are considered related, denoted as $i \sim j$. Notably, this relationship is symmetric and transitive, *i.e.,*

$$i \sim j \Leftrightarrow j \sim i, \tag{1}$$

$$i \sim j, j \sim k \Rightarrow i \sim k. \tag{2}$$

STL aims to map skeleton points from voxel space to topological metric space, ensuring that the distance between two points in topological space is proportional to their topological

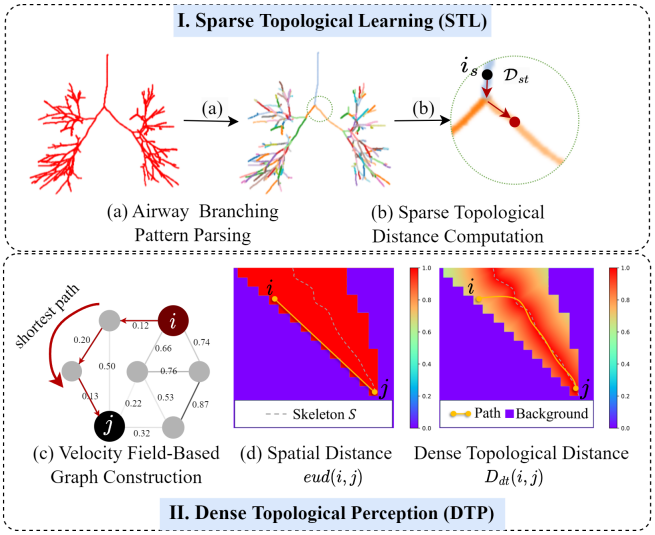

Fig. 2. Sparse topological learning and dense topological perception.

distance along the airway tree skeleton. Therefore, the sparse topology learning loss between point i and j $(i, j \in \Omega_{skel})$ is formulated as:

$$l_1 = \begin{cases} SmoothL_1\left[\hat{\mathcal{D}}_{st}(\hat{T}_i, \hat{T}_j) - \alpha_1 \mathcal{D}_{st}(i,j)\right], if\ i \sim j, \\ \beta * smoothL_1\left[\hat{\mathcal{D}}_{st}(\hat{T}_i, \hat{T}_j) - \tau\right], otherwise, \end{cases} \tag{3}$$

where $\hat{T}_i$ represents the embedding vector of the i^{th} voxel in the predicted topological metric space. $\hat{\mathcal{D}}_{st}(\hat{T}_i, \hat{T}_j) = \left\|\hat{T}_i - \hat{T}_j\right\|$, $\alpha_1 = 1/max(h,w,l)$ is the proportionality factor to normalize the $\mathcal{D}_{st}$. We set τ = 5 to maximize the topological space distance between uncorrelated voxel pairs. To ensure $l_1 \in (0,1)$, $\beta = 1/\tau = 1/5$. Thus, the total sparse topology learning loss is

$$\mathcal{L}_{stl} = \frac{1}{m} \sum_{i\in\Omega_{skel}} \sum_{j\in\Omega_{skel}} l_1(i,j), i \neq j, \tag{4}$$

where m is the total number of such voxel pairs.

2.2. Dense Topological Perception (DTP)

STL explicitly constructs the topological relationships at the skeleton level, but relying solely on individual points for topological preservation is insufficient. We argue that *in challenging regions for topological detection, such as branch breakages, every voxel is critical*. To address this, we aim to model local pixel relationships by defining dense topological metrics, thereby enhancing dense topological perception.

The question arises: how can we define the dense topological distance metric $\mathcal{D}_{dt}(i,j)$ between any two points i and j? We argue that this metric must satisfy the following

key attributes: 1) Topological distance refers to the length of the shortest topological path connecting two points, *i.e.* $\mathcal{D}_{dt}(i,j) := \min_{p \in \mathcal{P}_{i \to j}} [p]$, where $\mathcal{P}_{i \to j}$ is the set of all possible topological paths from i to j; 2) The topological path captures the branching trajectory / topological structure, with its length $\mathcal{D}_{dt}(i,j)$ being no less than the Euclidean distance $\mathcal{D}_{eud}(i,j)$; 3) The shortest topological path between two points is symmetric in both directions, *i.e.*, $p_{i \to j} \Leftrightarrow p_{j \to i}$.

Thus, the CT image is modeled as a 3D weighted graph $G(V,E)$, where each voxel i corresponds to an edge in the graph. The weight of each edge e_i is negatively correlated with the Euclidean distance from point i to the nearest skeleton point j, that is

$$e_i = 1/(\min_{j \in S}(\mathcal{D}_{eud}(i,j)) + 1). \tag{5}$$

Secondly, we identify a local hard region through hard skeleton sampling. Taking a hard skeleton point j as an example, the hard region is defined as the local neighborhood of j, *i.e.*, a spherical region $B_r(j) = \left\{x \in \mathbb{R}^3 \mid \|x - j\| \leq r\right\}$ with radius r centered at j. Finally, DTP utilizes dense connections within the hard region to enhance its local topological relationships. For any two points i and j $(i \neq j)$ within the neighborhood, the dense topological distance is defined as the shortest path length between the two points, calculated by a fast marching algorithm from scikit-fmm [20] with the weighted graph G as the velocity field, denoted as:

$$\mathcal{D}_{dt}(i,j) := \min_{p \in \mathcal{P}_{i \to j}} \int_0^1 \|\nabla G(x)(p(\xi)) \cdot \mathbf{u}(\xi)\| \, d\xi, \tag{6}$$

where $\mathbf{u}$ is a unit vector tangent to the direction of path p. As shown in Fig. 2II, compared to $\mathcal{D}_{eud}$, $\mathcal{D}_{dt}$ better captures the branching direction. The DTP loss of (i,j) is:

$$l_2 = \begin{cases} SmoothL_1(\mathcal{D}_{dt}(i,j) - \alpha_2 \hat{\mathcal{D}}_{dt}(\hat{T}_i, \hat{T}_j)), if\ i,j \in \Omega_f, \\ \beta * SmoothL_1\left[\hat{\mathcal{D}}_{dt}(\hat{T}_i, \hat{T}_j) - \tau\right], otherwise. \end{cases} \tag{7}$$

Similarly, $\hat{\mathcal{D}}_{dt}(\hat{T}_i, \hat{T}_j) = \left\|\hat{T}_i - \hat{T}_j\right\|$ and Ω_f denotes the foreground region. $\alpha_2 = 1/2r$ is the local normalization factor. The parameter τ is set to 5 to emphasize the inter-class differences between the foreground and background. To balance computational efficiency and segmentation performance, 1) During hard skeleton sampling, we consider both the prediction difficulty of skeleton points and the spatial distance between them. Specifically, we do adaptive skeleton point sparsification according to the airway depth, with deeper regions having a lower sparsification rate. This strategy avoids sampling overly close skeleton points, preventing excessive overlap in different difficult regions and saving computational resources. On this basis, DTP selects only the top k skeleton points with the highest prediction difficulty as sampling points for local topological learning in each iteration. 2) In each hard region, we randomly select n point pairs in each iteration to participate in the computation. The set of these point pairs is denoted as: $P_n(B_r(j)) = \{(x_k, x_l) \mid x_k, x_l \in B_r(j), k \neq l\}, \quad 1 \leq k,l \leq n$. The DTP loss is defined as follows:

$$\mathcal{L}_{dtp} = \frac{1}{k} \sum_{s \in S_k} \frac{1}{n} \sum_{(i,j) \in P_n(B_r(s))} l_2(i,j), \tag{8}$$

where $S_k \subset S$ represents the set of top k skeleton points with the lowest prediction probability.

2.3. Training of the ExpTopo framework

Anchored in the airway tree structure, STL and DTP strengthen voxel-level topology links through global sparsity and local density, alongside voxel-based airway predictions in training. Consequently, the overall objective function is formulated as:

$$\mathcal{L} = \mathcal{L}_{seg} + \gamma_1 \mathcal{L}_{stl} + \gamma_2 \mathcal{L}_{dtp}, \tag{9}$$

where $\mathcal{L}_{seg}$ is DiceFocal loss [21] imposed on the prediction $\hat{Y}$and Ground Truth Y. γ_1 and γ_2 are the trade-off weights.

3. EXPERIMENTS

Datasets and Implementation Details. We conduct our experiments on two public datasets: Binary Airway Segmentation (BAS) Dataset [6] consisting of 90 CT scans and ATM22 Dataset [22] consisting of 299 CT scans. All algorithms are developed and tested with a 4:1 train-test split. A modified 3D UNet architecture is selected as the backbone, each block in the encoder or decoder contains two convolutional layers followed by Instance Normalization. The preprocessing includes image value truncation, normalization and lung extraction. Sliding window is adopted to sample patches from each CT image with the size of $128 \times 128 \times 128$ with a half stride. Our models are trained by Adam optimizer with an initial learning rate of 0.01. The learning rate is divided by 10 in the 20^{th}, 40^{th}, and 60^{th} epoch. Results are thresholded by 0.5 for binarization and obtained by the largest connected component analysis. Experimentally, we set $\gamma_1 = \gamma_2 = 1, d = 8$ by default. Referring to [7], we adopt volumetric-based and topology-based metrics for evaluation. BD and TD mean branches detected and tree-length detected, respectively. Addtionally, BD^* means that a branch is considered detected only when 80% branch voxels are correctly classified.

Quantitative Results. Table 1 presents a comparison between ExpTopo and several other approaches, encompassing classical segmentation models [23, 24], popular airway tree segmentation algorithms from recent years [5–8], dense connectivity modeling methods [6, 25], and centerline-based topological enhancement strategies [14, 16]. Specifically, AirwayNet [6] and CoANet [25] employ connectivity prediction and connectivity modeling, respectively, to achieve connectivity-aware tubular structure segmentation either directly or indirectly. As typical examples of sparse topology

Table 1. Quantitative Results on the BAS Dataset.

Methods	BD*(%) ↑	BD(%) ↑	TD(%)↑	TPR(%) ↑	FPR(%) ↓	DSC(%) ↑
U-Net3D [23]*	$62.04_{\pm 23.6}$	$77.82_{\pm 21.6}$	$73.48_{\pm 22.2}$	$86.94_{\pm 14.3}$	$.010_{\pm .006}$	$90.58_{\pm 9.5}$
U-Net3D [23]†	$69.33_{\pm 21.1}$	$82.70_{\pm 16.2}$	$80.38_{\pm 17.4}$	$90.86_{\pm 8.8}$	$.017_{\pm .008}$	$92.05_{\pm 4.2}$
SwinUnetr [24]†	$75.35_{\pm 17.1}$	$86.03_{\pm 10.9}$	$84.82_{\pm 12.5}$	$91.35_{\pm 7.9}$	$.023_{\pm .012}$	$91.73_{\pm 3.8}$
Navi-Airway [5]	$73.37_{\pm 20.8}$	$85.87_{\pm 15.1}$	$82.32_{\pm 16.9}$	$91.81_{\pm 9.8}$	$.057_{\pm .015}$	$86.48_{\pm 5.5}$
Qin *et al* [7]	$69.04_{\pm 21.3}$	$81.93_{\pm 16.9}$	$79.09_{\pm 18.4}$	$90.34_{\pm 9.3}$	$.022_{\pm .007}$	$92.11_{\pm 4.7}$
Wingnet [8]	$74.40_{\pm 19.7}$	$85.14_{\pm 13.7}$	$84.04_{\pm 14.7}$	$90.25_{\pm 8.0}$	$.016_{\pm .009}$	$92.08_{\pm 3.8}$
AirwayNet [6]	$69.20_{\pm 18.8}$	$84.23_{\pm 12.9}$	$82.10_{\pm 14.5}$	$87.20_{\pm 7.7}$	$.011_{\pm .009}$	$90.76_{\pm 3.5}$
CoANet [25]	**75.83**$_{\pm 19.2}$	$86.13_{\pm 13.4}$	$84.82_{\pm 15.1}$	**92.31**$_{\pm 7.9}$	$.022_{\pm .009}$	$92.14_{\pm 3.5}$
clDice [16]	$63.53_{\pm 22.5}$	$78.76_{\pm 18.8}$	$74.90_{\pm 19.7}$	$87.75_{\pm 11.2}$	**.010**$_{\pm .005}$	$91.29_{\pm 6.3}$
clCE [14]	$63.67_{\pm 22.1}$	$79.10_{\pm 18.8}$	$75.41_{\pm 19.4}$	$88.31_{\pm 10.9}$	$.012_{\pm .007}$	$91.30_{\pm 5.9}$
Ours	$75.79_{\pm 18.6}$	**87.26**$_{\pm 11.8}$	**85.25**$_{\pm 13.9}$	$92.27_{\pm 7.3}$	$.021_{\pm .009}$	**92.33**$_{\pm 3.2}$

Note: * indicates that the objective function during training is the Dice loss. † indicates that DiceFocalLoss is used by default.

Table 2. Quantitative Results on the ATM22 Dataset.

Methods	BD*(%) ↑	BD(%) ↑	TD(%)↑	TPR(%) ↑	FPR(%) ↓	DSC(%) ↑
U-Net3D [23]	$83.03_{\pm 12.4}$	$96.26_{\pm 5.0}$	$88.86_{\pm 8.9}$	$94.94_{\pm 3.8}$	**.019**$_{\pm .014}$	**94.26**$_{\pm 2.2}$
SwinUnetr [24]	$84.32_{\pm 12.7}$	$95.37_{\pm 7.9}$	$89.39_{\pm 10.0}$	$93.80_{\pm 6.6}$	$.021_{\pm .013}$	$93.43_{\pm 4.0}$
Qin *et al* [7]	$83.80_{\pm 14.0}$	$95.49_{\pm 8.4}$	$89.14_{\pm 11.0}$	$95.11_{\pm 10.0}$	$.021_{\pm .012}$	$93.98_{\pm 2.7}$
Wingnet [8]	$87.21_{\pm 11.6}$	**97.99**$_{\pm 1.8}$	$91.59_{\pm 6.9}$	$94.67_{\pm 3.8}$	$.031_{\pm .035}$	$92.66_{\pm 4.2}$
CoANet [25]	$87.11_{\pm 14.0}$	$96.83_{\pm 5.6}$	$91.51_{\pm 8.6}$	$95.31_{\pm 4.2}$	$.022_{\pm .012}$	$94.02_{\pm 2.7}$
Ours	**88.03**$_{\pm 10.9}$	$97.44_{\pm 3.5}$	**92.15**$_{\pm 7.6}$	**95.95**$_{\pm 3.1}$	$.023_{\pm .013}$	$94.20_{\pm 2.1}$

Table 3. Results of ablation study on BAS dataset.

Manner		r	n	BD(%) ↑	TD(%)↑	TPR(%) ↑	FPR(%) ↓	DSC(%) ↑
	-	-	-	$82.70_{\pm 16.2}$	$80.38_{\pm 17.4}$	$90.86_{\pm 8.8}$	**.017**$_{\pm .008}$	$92.05_{\pm 3.8}$
	DTP	3	all	$84.92_{\pm 15.0}$	$83.02_{\pm 16.1}$	$91.77_{\pm 7.8}$	$.024_{\pm .011}$	$92.10_{\pm 4.0}$
Multi Branch	DTP	8	300	$85.17_{\pm 14.7}$	$83.43_{\pm 15.7}$	$91.80_{\pm 7.6}$	$.020_{\pm .009}$	$92.12_{\pm 4.7}$
	DTP	8	800	$85.67_{\pm 14.2}$	$83.95_{\pm 15.4}$	$91.87_{\pm 7.3}$	$.021_{\pm .008}$	$92.32_{\pm 4.4}$
	DTP	8	all	$86.70_{\pm 13.5}$	**85.30**$_{\pm 13.9}$	**92.35**$_{\pm 6.8}$	$.026_{\pm .009}$	$92.22_{\pm 3.5}$
	DTP	3	all	$85.12_{\pm 15.2}$	$83.13_{\pm 16.3}$	$91.80_{\pm 8.0}$	$.022_{\pm .010}$	$92.19_{\pm 3.8}$
Multi Head	DTP	8	300	$85.20_{\pm 14.8}$	$83.49_{\pm 15.9}$	$91.92_{\pm 7.9}$	$.023_{\pm .011}$	$92.24_{\pm 4.0}$
	DTP	8	800	$85.78_{\pm 14.6}$	$84.10_{\pm 15.6}$	$92.09_{\pm 7.7}$	$.023_{\pm .010}$	$92.25_{\pm 4.1}$
	DTP	8	all	$86.43_{\pm 13.1}$	$85.11_{\pm 14.5}$	$92.20_{\pm 7.5}$	$.025_{\pm .008}$	**92.35**$_{\pm 3.5}$
	DTP+STL	8	all	**87.26**$_{\pm 11.8}$	$85.25_{\pm 13.9}$	$92.27_{\pm 7.3}$	$.021_{\pm .009}$	$92.33_{\pm 3.2}$

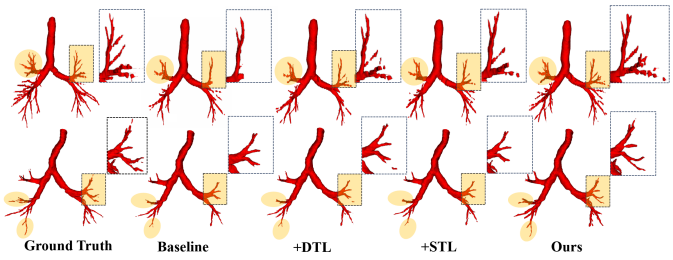

Fig. 3. Comparison of Segmentation results.

learning, clDice [16] and clCE [14] introduce clDice and clCE loss upon the basic dice loss, respectively, to enhance detection accuracy at the centerline level. From Table 1, we observe that **1)** ExpTopo achieved the best performance in topological-level metrics BD, TD, and overall voxel-level metric DSC, highlighting the significant advantage of incorporating global sparse topological representation and local dense topological representation in voxel-level classification for tubular structures. **2)** clDice and clCE improve upon their baseline U-Net3D* (U-Net with Dice loss), but underperform compared to U-Net3D† (U-Net with DiceFocal loss). This may be due to Focal loss's ability to address class imbalance by uniformly emphasizing hard-to-classify regions, rather than individual centerline voxels. This reinforces the core argument of the paper, that *each voxel in difficult regions is essential for preserving both voxel-wise and topological accuracy*. **3)** When implemented on the larger ATM22 dataset, as shown in Table 2, the algorithm consistently maintains competitive performance even without any backbone architecture optimization or post-processing.

Ablation Study. We conduct ablation studies to explore two designs of the algorithm: the method of integrating topological learning into the network architecture (multi-branch *vs.* multi-head) and the introduction of DTP and STL. As illustrated in Table 3, we can observe: **1)** Whether using a multi-head or multi-branch strategy, the inclusion of the topological learning branch resulted in comprehensive improvements, demonstrating the importance and reliability of explicit topological learning. In comparison, the multi-branch approach showed slightly better improvement in topology-level metrics (BD 86.70 *v.s.* 86.43), while the multi-head approach achieved the best overall DSC score (DSC 92.35 *v.s.* 92.22) with fewer added parameters (+2.73% *v.s.* +58.3%). **2)** The introduction of both DTP and STL in the multi-head mode led to significant performance improvements compared to the baseline, particularly at the topological level (BD +4.56%, TD +4.87%). In addition, we conducted a detailed investigation into the radius r of the hard region and the number of point pairs n sampled within the region in dense topological learning. We found that as the radius r increased and the number of sampled point pairs grew, the effectiveness of dense topological learning became more pronounced, which aligns with the motivation of the algorithm. In difficult regions, the denser the topological network, the better the relational learning, thereby making it easier to compensate for or correct weakly detected distant areas.

Qualitative Results. As shown in Fig. 3, both DTP and STL notably improve the accurate identification of distal fine branches. The combined application of both strengthens the model's topological awareness from both a global sparse and local dense perspective, ultimately achieving optimal segmentation performance.

4. CONCLUSION

This paper proposes an approach that enhances the model's perception of topological structures by explicitly incorporating effective topological metric learning on top of voxel-level semantic segmentation. Specifically, it optimizes topological integrity and addresses local detection deficiencies by introducing global sparse and local dense topological learning. Extensive experiments demonstrate the superiority of our algorithm, with significant improvements in topological accuracy compared to baseline models.

5. ACKNOWLEDGMENTS

The study was supported by Natural Science Foundation of China under Grant 62271465, Suzhou Basic Research Program under Grant SYG202338, and Open Fund Project of Guangdong Academy of Medical Sciences, China (No. YKY-KF202206).

6. REFERENCES

[1] S. A. Christenson et al, "Chronic obstructive pulmonary disease," *The Lancet*, vol. 399, no. 10342, pp. 2227–2242, 2022.

[2] J. Guo et al, "Coarse-to-fine airway segmentation using multi information fusion network and cnn-based region growing," *Computer Methods and Programs in Biomedicine*, vol. 215, pp. 106610, 2022.

[3] M. Zhang et al, "Fda: Feature decomposition and aggregation for robust airway segmentation," in *MICCAI*, 2021, pp. 25–34.

[4] M. Zhang et al, "Towards connectivity-aware pulmonary airway segmentation," *IEEE JBHI*, 2023.

[5] A. Wang et al, "NaviAirway: a Bronchiole-sensitive Deep Learning-based Airway Segmentation Pipeline," 2022.

[6] Y. Qin et al, "AirwayNet-SE: A Simple-Yet-Effective Approach to Improve Airway Segmentation Using Context Scale Fusion," in *IEEE ISBI*, Apr. 2020, pp. 809–813.

[7] Y. Qin et al, "Learning Tubule-Sensitive CNNs for Pulmonary Airway and Artery-Vein Segmentation in CT," *IEEE TMI*, vol. 40, no. 6, pp. 1603–1617, 2021.

[8] H. Zheng et al, "Alleviating Class-Wise Gradient Imbalance for Pulmonary Airway Segmentation," *IEEE TMI*, , no. 9, pp. 2452–2462, 2021.

[9] M. Zhao et al, "Gdds: Pulmonary bronchioles segmentation with group deep dense supervision," 2023.

[10] H. Zheng et al, "Refined Local-imbalance-based Weight for Airway Segmentation in CT," in *MICCAI*, pp. 410–419. 2021.

[11] Y. Wu et al, "Two-stage contextual transformer-based convolutional neural network for airway extraction from ct images," *Artificial Intelligence in Medicine*, vol. 143, pp. 102637, 2023.

[12] Y. Qin et al, "Learning Bronchiole-Sensitive Airway Segmentation CNNs by Feature Recalibration and Attention Distillation," in *MICCAI*, vol. 12261, pp. 221–231. 2020.

[13] Zeyu Tang, Yang Nan, Simon Walsh, and Guang Yang, "Adversarial transformer for repairing human airway segmentation," *IEEE JBHI*, 2023.

[14] C. Acebes et al, "The centerline-cross entropy loss for vessel-like structure segmentation: Better topology consistency without sacrificing accuracy," in *MICCAI*. Springer, 2024, pp. 710–720.

[15] D. Wang et al, "PointScatter: Point Set Representation for Tubular Structure Extraction," in *ECCV*, 2022, pp. 366–383.

[16] S. Shit et al, "cldice-a novel topology-preserving loss function for tubular structure segmentation," in *CVPR*, 2021, pp. 16560–16569.

[17] Xiaoling Hu, "Structure-aware image segmentation with homotopy warping," in *NeurIPS*, 2022.

[18] X. Zhang et al, "Progressive deep segmentation of coronary artery via hierarchical topology learning," in *MICCAI*, 2022, pp. 391–400.

[19] Deepak Keshwani, Yoshiro Kitamura, Satoshi Ihara, Satoshi Iizuka, and Edgar Simo-Serra, "Topnet: Topology preserving metric learning for vessel tree reconstruction and labelling," in *MICCAI*. Springer, 2020, pp. 14–23.

[20] J Furtney, "Scikit-fmm software," 2015.

[21] T.-Y. Ross et al, "Focal loss for dense object detection," in *CVPR*, 2017, pp. 2980–2988.

[22] M. Zhang et al, "Multi-site, multi-domain airway tree modeling," *Medical Image Analysis*, vol. 90, pp. 102957, 2023.

[23] Özgün. Çiçek et al, "3d u-net: Learning dense volumetric segmentation from sparse annotation," in *MICCAI*, 2016, pp. 424–432.

[24] A. Hatamizadeh et al, "Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images," in *International MICCAI Brainlesion Workshop*, 2021, pp. 272–284.

[25] J. Mei et al, "Coanet: Connectivity attention network for road extraction from satellite imagery," *IEEE TIP*, vol. 30, pp. 8540–8552, 2021.