BCNet: Bronchus Classification via Structure Guided Representation Learning

Wenhao Huang, Haifan Gong, *Graduate Student Member, IEEE*, Huan Zhang, Yu Wang, Xiang Wan, Guanbin Li, *Member, IEEE*, Haofeng Li, *Member, IEEE*, and Hong Shen

***Abstract*— CT-based bronchial tree analysis is a key step for the diagnosis of lung and airway diseases. However, the topology of bronchial trees varies across individuals, which presents a challenge to the automatic bronchus classification. To solve this issue, we propose the Bronchus Classification Network (BCNet), a structure-guided framework that exploits the segment-level topological information using point clouds to learn the voxel-level features. BCNet has two branches, a Point-Voxel Graph Neural Network (PV-GNN) for segment classification, and a Convolutional Neural Network (CNN) for voxel labeling. The two branches are simultaneously trained to learn topology-aware features for their shared backbone while it is feasible to run only the CNN branch for the inference. Therefore, BCNet maintains the same inference efficiency as its CNN baseline. Experimental results show that BCNet significantly exceeds the state-of-the-art methods by over 8.0% both on F1-score for classifying bronchus. Furthermore, we contribute BronAtlas: an open-access benchmark of bronchus imaging analysis with high-quality voxel-wise annotations of both anatomical and abnormal bronchial segments. The benchmark is available at https://osf.io/pskr9/?viewonly=94fa3d87274b4095ac9a4b88cc9a1341.**

***Index Terms*— Bronchus classification, benchmark, CT imaging, guidance-based distillation, graph neural network, convolution neural network, multi-task learning, point cloud.**

Manuscript received 21 May 2024; revised 27 June 2024; accepted 14 August 2024. Date of publication 23 August 2024; date of current version 2 January 2025. This work was supported in part by Shenzhen Science and Technology Program under Grant JCYJ20220818103001002; in part by Guangdong Provincial Key Laboratory of Big Data Computing, The Chinese University of Hong Kong, Shenzhen; in part by Longgang District Special Funds for Science and Technology Innovation under Grant LGKCSDPT2023002; in part by the National Natural Science Foundation of China under Grant 62102267 and Grant 62322608; in part by the Fundamental Research Funds for the Central Universities under Grant 22lgqb25; and in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2023A1515011464. *(Wenhao Huang and Haifan Gong contributed equally to this work.) (Corresponding authors: Guanbin Li; Haofeng Li.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Review Committee of Peking University People's Hospital under Application No. 2021PHA038-001.

Wenhao Huang, Huan Zhang, Yu Wang, and Hong Shen are with InferVision, Beijing 100000, China (e-mail: hwenhao@infervision.com; zhuan@infervision.com; wyua@infervision.com; shong@infervision.com).

Haifan Gong is with Shenzhen Research Institute of Big Data, Shenzhen 518172, China, and also with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen 518172, China (e-mail: haifangong@link.cuhk.edu.cn).

Xiang Wan and Haofeng Li are with Shenzhen Research Institute of Big Data, Shenzhen 518172, China (e-mail: wanxiang@sribd.cn; lhaof@sribd.cn).

Guanbin Li is with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China (e-mail: liguanbin@mail.sysu.edu.cn).

Digital Object Identifier 10.1109/TMI.2024.3448468

I. Introduction

THE lung airway analysis based on CT imaging is clinically critical since it provides quantitative information to assist lung disease diagnosis and surgical navigation [1], [2], [3]. As the basis of lung airway analysis, bronchial tree reconstruction usually comprises two steps [4], [5], [6]. The first step is to extract the binary airway-tree mask from CT imaging. The second step is to label major anatomical branches via bronchus segment classification. The automatic bronchial tree reconstruction can support more clinical processes, such as individual airway tree phenotype matching and lung lobe or lung segment classification [7], [8]. Moreover, accurate labeling of bronchial segments is helpful for clinicians to understand which anatomic segments are affected by disease and can be helpful for surgical planning [9], [10], [11]. In the automatic bronchus segment classification, there still exists a challenge: how to model and exploit the topology of the bronchial tree efficiently.

Motivated by the importance of structure priors in medical image analysis [12], [13], we developed a Bronchus Classification Network (BCNet), a framework designed to effectively classify bronchus segmentation from CT images. BCNet leverages the inherent topology of the bronchial tree as structural guidance by incorporating a Point-Voxel Graph Neural Network (PV-GNN) to enhance the representation learning of a Convolutional Neural Network (CNN). The core idea behind BCNet is to utilize segment-wise topological information (via the GNN) to guide the learning of voxel-wise convolution features (via the CNN). This approach enables the CNN branch to better distinguish bronchus segments at the voxel level, as illustrated in Figure 1. Notably, BCNet can operate solely with the convolution branch during inference, eliminating the need for the graph model and thereby achieving the accuracy benefits of GNN training while maintaining the computational efficiency of the CNN baseline.

To construct the GNN branch of BCNet, we integrate high-dimensional local features from the CNN with positional and

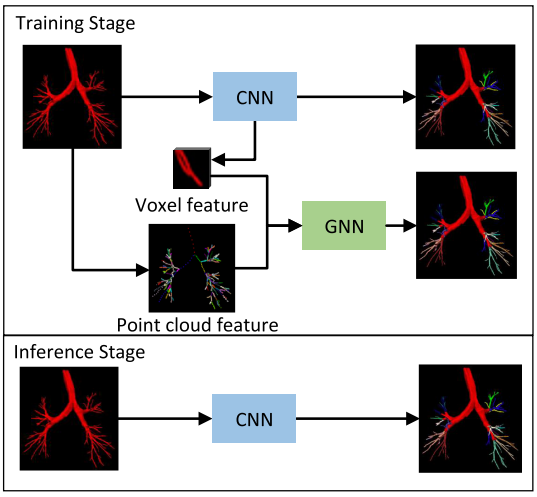

Fig. 1. The key idea of our BCNet. The upper part shows the training process involves the GNN and CNN. The GNN and CNN parts share the same voxel-wise CNN feature map, which is used to represent the bronchus in the GNN part. By computing loss in the GNN part, we encourage the shared feature map to better describe the bronchus segments and the topological relationships between segments. It can be viewed as additional contextual guidance that is different from and complements the voxel-level supervision in the CNN part. The bottom part shows that our inference phrase only needs the CNN part.

angular information described by point clouds. Recognizing that adjacent bronchus segments often belong to the same category, we introduce a Neighborhood Consistency Regularization as a loss term to enhance model training. For evaluation, we manually annotated airway branches in 100 CT scans, comprising 60 scans from public datasets and 40 from a collaborating hospital. This comprehensive annotation effort ensures robust assessment of both anatomical and abnormal labels. Overall, our contributions are summarized as follows:

- BronAtlas, a benchmark containing 100 bronchial cases with accurate voxel-level segmentation masks and anatomical categories including the abnormal bronchus segment;
- BCNet, a structure-guided representation learning framework to classify the bronchus from lung CT imaging, which significantly outperforms the state-of-the-art methods on BronAtlas;
- PV-GNN, a Point-Voxel Graph Neural Network to perform the segment-level bronchus classification, which uses the point cloud and convolutional feature to represent the relative position and semantics of the bronchus segment, respectively. By taking both the convolution feature and point cloud feature as the node feature to construct the graph, the PV-GNN has shown its superiority through ablation studies.

II. Related Work

A. Bronchus Analysis

In recent years, deep learning (DL) has shown its advance for bronchus segmentation [4], [14]. Single-stage networks like U-Net [15], 2.5D Net [5], [16], 3D U-Net [2], [6], [17], [18], [19], [20], [21], [22], [23] have been employed to segment the bronchus. Meng et al. [6] proposed an image tracking method for the segmentation task. Qin et al. [19], [20] proposed a CNNs-based airway segmentation method which works better on superior sensitivity to tenuous peripheral bronchioles. Zheng et al. [22] adopted gradient erosion and dilation operators to alleviate the inter-class imbalance issue in segmenting bronchus. Zheng et al. [24] adjusted the gradient ratio of each airway point based on the quantification of local class imbalance, and further proposed a weight enhancement strategy to handle the hard-to-segment regions. Tang et al. [25] propose the adversarial transformer for airway segmentation. Nan et al. [26] propose the fuzzy attention neural network to tackle discontinuity in airway segmentation.

Bronchus classification [4], [27], [28] is a challenging task due to that the topology of bronchial trees substantially varies across individuals. It is usually based on the results of bronchus segmentation methods [2], [16], [17], [18], [19], [20], [21], [22], [24], [25]. The conventional methods for bronchus classification include geodesic matching [28] and probabilistic hypergraph matching [29]. Wang et al. [8] achieved lobar-level bronchus classification using key-point detection. Lo et al. [30] proposed a bottom-up approach for bronchial labeling based on rules. For deep learning-based methods, Zhao et al. [31] analyzed the average inclination angle of bronchial segments in the training set and applied linear programming to post-process the airway structure predicted by neural models. Nadeem et al. [7] proposed a neural network of two stages that label the lobar-level and segment-level bronchus, respectively. Yu et al. [32] propose the tree neural network for airway labeling.

B. Graph Models for Segmentation

Recently, Graph Neural Networks (GNNs) [33], [34], [35] have become increasingly important in the biomedical domain [36], [37], [38], [39], [40], particularly for addressing segmentation tasks [41], [42], [43], [44], [45], [46]. For instance, Garcia-Uceda et al. [41] enhanced airway binary segmentation by replacing the deepest convolutional layer in a U-Net [15] with graph convolutions [33], demonstrating the potential of GNNs in improving segmentation results. Similarly, Selvan et al. [42] introduced a GCN-based [33] mean-field network to refine segmentation outputs produced by a 3D-UNet [15], further highlighting the role of GNNs in segmentation refinement. Tan et al. [44] approached bronchial segmentation and classification as a semantic segmentation problem, proposing a two-stage framework specifically designed for bronchus classification. Zhao et al. [43] developed a prototype-based graph neural network aimed at detecting abnormal bronchus structures, showcasing the application of GNNs in identifying pathological changes.

Therefore, there is a pressing need for a more efficient and effective framework that leverages voxel-wise features and point-cloud topology to guide voxel-level representation learning for bronchus classification. Our proposed Bronchus Classification Network (BCNet) addresses this gap by integrating the structural topology of the bronchial tree into the learning process. BCNet employs a Point-Voxel Graph Neural Network (PV-GNN) to enhance the representation learning capabilities of a Convolutional Neural Network (CNN).

III. BronAtlas: A Benchmark for Bronchus Classification

A. Motivation

The category and the abnormality of bronchus segments play an important role in computer-aided bronchial disease diagnosis. However, the existing bronchial classification dataset does not contain the categorical label of abnormal bronchial segments. Moreover, we observe that the existing bronchial segmentation datasets [2], [47], [48] do not provide the category of bronchial segments. To facilitate the development of automatic bronchus diagnosis, we contribute a new benchmark, BronAtlas, for bronchus segmentation and classification. BronAtlas contains 100 cases of lung CT imaging with voxel-level annotations of 20 segmental bronchi categories (i.e., 10 from the right lung, 8 from the left lung, 1 for the main trachea, and 1 for abnormal bronchus). We will make this dataset available after acceptance.

B. Sample Collection and Annotation

We collect 60 cases from the currently available databases EXACT'09 [47] and LIDC [48]. The remaining 40 cases are collected from our cooperative hospital. For the currently public samples and the newly collected ones, each CT scan is annotated by two experts in a two-step annotation process. The experts first annotate binary masks for the airway segmentation and then label 20 segmental bronchi at the voxel level according to the anatomical structure which is shown in Fig. 2. With the binary segmentation of the bronchus, the radiologists label the 18 segments of the bronchus and the main trachea according to [27], where 10 of these segmental bronchi are from the right lung (i.e., RB1 to RB10) while the remaining 8 segmental bronchi are from the left lung (i.e., LB1 + 2, LB3,..., LB7 + 8, LB9, LB10). During the annotation process, the radiologists make sure that except for the abnormal branches, the bronchial segments should be connected to the main bronchial trunk, and that the bronchial segments do not have any problems of misclassification, under classification, non-abnormal co-drying, or cross-color. For the bronchus segments that do not belong to the above classes, the radiologists regard them as abnormal bronchus. The process of data acquisition and investigation follows the principles outlined in the declaration of Helsinki [49], and has received the appropriate approvals from the institutional ethical committee.

C. Dataset Summary

For the development of automatic bronchus diagnosis, we contribute a new benchmark, BronAtlas, for bronchus classification. BronAtlas has 100 cases of lung CT imaging with voxel-level annotations of 20 segmental bronchi categories (*i.e.* , 10 from the right lung, 8 from the left lung, 1 for the main trachea, and 1 for abnormal bronchus). The "abnormal bronchus segment" denotes the congenital abnormal bronchus. There are 33 abnormal cases in the training and validation set and 12 abnormal cases in the test set. The visualization of the 20 bronchial categories is in Figure 2. The details of the samples in the benchmark are in Table I. We mix these data from different sources together and split the whole dataset into a training/validation/test set of 63/7/30 cases, respectively. The patient data used in this study was anonymized, and all personally identifiable information was removed before the model training and evaluation processes. This procedure ensures the privacy of the individuals whose CT scans were included in the study. We attempted to minimize bias by using a diverse dataset, including CT scans from individuals of different ages, genders, and health conditions. It would be our future work to construct a more systematic approach to addressing bias, such as stratified sampling or bias audits.

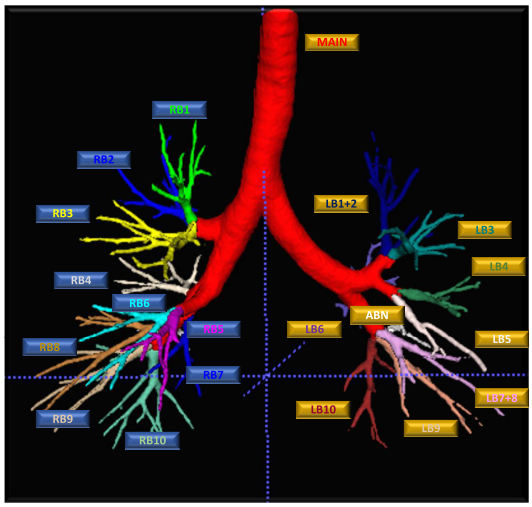

Fig. 2. The visualization results of the bronchus segment labels. The segmental bronchi from the right lung (i.e., RB1 to RB10) and segmental bronchi from the left lung (i.e., LB1 + 2, LB3,..., LB7 + 8, LB9, LB10) are defined according to [27]. "MAIN" denotes the main trachea while "ABN" denotes the abnormal bronchus segment. The radiologists use the binary segmentation of the bronchus and make sure the bronchial segments are connected to the main bronchial trunk except for the abnormal branches. For the bronchus segments that do not belong to the above classes, the radiologists regard them as abnormal bronchus.

TABLE I
DETAILS OF THE BRONATLAS: A BENCHMARK FOR BRONCHUS CLASSIFICATION WITH A VOXEL-WISE MASK. WE COLLECT 60 CASES FROM THE CURRENTLY AVAILABLE DATABASES EXACT'09 [47] AND LIDC [48]. THE REMAINING 40 CASES ARE COLLECTED FROM OUR COOPERATIVE HOSPITAL. FOR THE CURRENTLY PUBLIC SAMPLES AND THE NEWLY COLLECTED ONES, EACH CT SCAN IS ANNOTATED BY TWO EXPERTS IN A TWO-STEP ANNOTATION PROCESS. THE EXPERTS FIRST ANNOTATE BINARY MASKS FOR THE AIRWAY SEGMENTATION, AND THEN LABEL 20 SEGMENTAL BRONCHI AT THE VOXEL LEVEL ACCORDING TO THE ANATOMICAL STRUCTURE IN FIGURE 2

Attributes of BronAtlas	Values
Resolution of axial slices	512×512
Thickness of axial slices	$0.55 \sim 1.00$ mm
#Samples in BronAtlas	100
#Samples from LIDC	40
#Samples from EXACT'09	20

IV. Method

To use the inherent topology of the bronchus to guide the representation learning for bronchus classification, we proposed the Bronchus Classification Network (BCNet), which

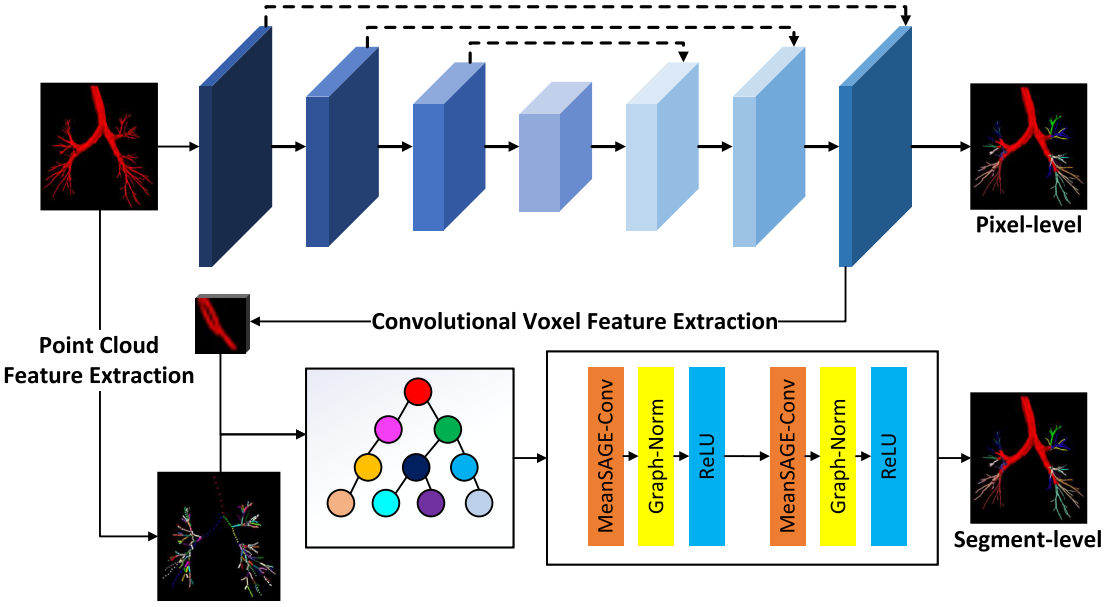

Fig. 3. Overview of the Bronchus Classification Network (BCNet). The upper part is a UNet that predicts the voxel-level categorical labels of the bronchus based on a binary bronchus mask. The lower part is a graph neural network (GNN) that predicts the categories of the bronchus at the segment level. In the training, the gradients of the GNN part are passed backward through the UNet to guide its representation learning.

is shown in Fig. 3. The proposed BCNet includes two components: a UNet for voxel-level bronchus labeling, and a Point-Voxel Graph Neural Network (PV-GNN) for segment-level bronchus labeling. The structure-guided representation learning in BCNet means that the UNet can benefit from the gradients propagated from the PV-GNN branch since the PV-GNN adopts the convolutional feature from the UNet branch to construct graph nodes.

A. Motivation

We employ the UNet to obtain the voxel-level bronchus segmentation result since the UNet can compute the high-dimension voxel feature for the following segment-level bronchus labeling. The motivation behind our proposed GNN is that the relative positional information represented by the point cloud can help model the intrinsic bronchus topology during the classification process. In the meanwhile, the high-dimension voxel feature could be a strong supplement for bronchus classification since it can implicitly capture other kinds of bronchus attributes such as the diameter of bronchi. Finally, we utilize the GNN to guide the learning process of the UNet through the convolution feature and improve the classification performance of the UNet. The joint learning of the GNN and CNN branches can be seen as a form of regularization. By encouraging the model to learn representations on two tasks, the model can achieve better generalization performance. Note that the gradients of GNN loss are passed through the CNN, while those of CNN loss are not sent into the GNN. Thus, the CNN is trained with the voxel-level loss and an additional segment-level loss via the feature construction procedure of GNN, while the GNN part is only trained with the segment-level loss. The CNN can learn from both tasks at the voxel level and segment level and exploit the shared information to improve performance without introducing additional parameters.

B. UNet for Voxel-Level Bronchus Labeling

We classify the bronchus at voxel level with a UNet [18] that is shown in the upper part of Fig. 3. The UNet can produce both the coarse bronchial prediction and the voxel-level feature for further segment-level classification. The UNet is trained with the cross-entropy loss, which is formulated as:

$$L_{unet} = L_{CE}(y_{pred}, y_{gt}), \tag{1}$$

where the ground truth and the prediction are denoted as y_{gt} and y_{pred}, respectively. It is worth noting that the background class is not considered.

C. Point-Voxel GNN for Segment-Level Bronchus Labeling

1) Definition of Point-Voxel Graph: Given a binary segmentation mask of the bronchus, we update the mask as its maximum connected region and define a point-voxel graph with the region, using mixed voxel and point-level features as node embeddings. The point-voxel graph was built, as shown in Fig. 4. First, we skeletonize the binary mask by extracting the centerline (see Fig. 4(b)) via an existing algorithm [50]. Second, according to the number N of foreground voxels in the 26-connected neighborhood of each voxel on the centerline, we define the endpoints ($N = 1$), edge points ($N = 2$), and division points ($N \geq 3$), as shown in Fig. 4(c). Finally, as Fig. 4(d) displays, we divide the branches into segments with these points. Each bronchial segment corresponds to a node in the proposed point-voxel graph. An edge of the graph is defined by the connectivity between two line segments that are divided by a division point. The details of our network structure are shown in Table II.

2) Point-Wise Coordinate Feature: We generate the point cloud from the bronchus mask to obtain the point-wise coordinate feature. We crop out the bounding box of each bronchial

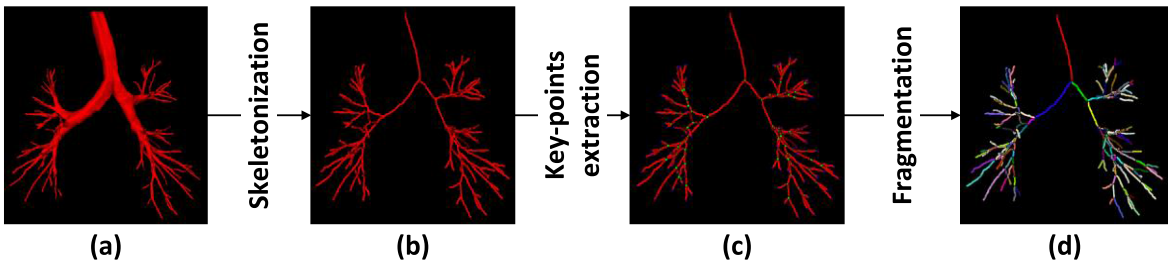

Fig. 4. Construction of bronchial tree: (1) skeletonize the bronchus; (2) extract the key points separating the bronchial fragments; (3) separate the bronchus into segments.

TABLE II
DETAILS OF THE UNET AND POINT-VOXEL GRAPH NEURAL NETWORK (PV-GNN) STRUCTURES AND PARAMETERS. "BN" DENOTES BATCH NORMALIZATION, "SK" DENOTES SKIP CONNECTION, "FC" DENOTES FULLY CONNECTED LAYER, "CONV" DENOTES CONVOLUTION, "MSC" DENOTES MEAN SAGE-CONVOLUTION, "GN" DENOTES GRAPH NORMALIZATION, AND "CONV-NORM BLOCK" COMBINES MSC, GN, AND RELU LAYERS. "K" DENOTES THE NUMBER OF POINTS PER NODE IN THE GRAPH

Module	Input Size/Channels	Network Layer/Operation	In/Out Channels	Description
		UNet Architecture		
Encoder	$128 \times 128 \times 128$	Stem Conv	1 / 24	Conv3d+BN+ReLU
	$64 \times 64 \times 64$	Residual block 1	24 / 48	Multiple layers with SK
	$32 \times 32 \times 32$	Residual block 2	48 / 96	Multiple layers with SK
	$16 \times 16 \times 16$	Residual block 3	96 / 192	Multiple layers with SK
Decoder	$32 \times 32 \times 32$	Upsample layer	192 / 192	Trilinear upscaling
	$32 \times 32 \times 32$	Double Conv	192+96 / 96	Conv3d+BN+ReLU
	$64 \times 64 \times 64$	Upsample layer	96 / 96	Trilinear upscaling
	$64 \times 64 \times 64$	Double Conv	96+48 / 48	Conv3d+BN+ReLU
	$128 \times 128 \times 128$	Upsample layer	48 / 48	Trilinear upscaling
	$128 \times 128 \times 128$	Double Conv	48+24 / 24	Conv3d+BN+ReLU
	$128 \times 128 \times 128$	FC layer	24 / 20	Reduce dimensions
		PV-GNN Architecture		
PV-GNN	$(24+3) \times K$	GNN Block 1	$(24+3) \times K$ / $(24+3) \times K$	Conv-Norm Block
	$(24+3) \times K$	GNN Block 2	$(24+3) \times K$ / $(24+3) \times K$	Conv-Norm Block
	$(24+3) \times K$	GNN Block 3	$(24+3) \times K$ / $(24+3) \times K$	Conv-Norm Block
	$(24+3) \times K$	Fully-connected layer	$(24+3) \times K$ / 20	Reduce dimensions

segment (i.e., a node in the graph) from the bronchus tree. Each segment is composed of the centerline voxels that are within the bounding box. Then the coordinate (X, Y, and Z) of each voxel is normalized to [0, 1] concerning the shape of the bounding box. Such normalization is expected to help describe the pose and angle of a segment. Let L be the length of the center line, We divide the centerline into $K-1$ parts with the same interval $\frac{L}{K-1}$, and collect K points from the start and the end points of these $K-1$ parts. For each bronchial segment, we define its *point-wise coordinate feature* as the list of three-dimensional coordinates of these K points. In this way, for each bronchial segment, we obtain its point cloud feature that contains three-dimensional coordinates of K voxels and has a length of $3 \times K$. The detailed analysis of the hyper-parameter K is shown in Table VII.

3) Voxel-Wise Convolution Feature: To obtain voxel features, we train a UNet (the upper part in Fig. 3) to predict the category of each bronchus branch. A 3D feature map is produced from the penultimate layer of the UNet. For each bronchial fragment, we locate the K voxels that are nearest to the above-mentioned K points and extract the convolution features from the feature map according to each voxel's indices. The *voxel-wise convolution feature* of the bronchial fragment is defined as the combination of the K features and has a length of $C \cdot K$, where C is the channel number of the last but one feature map ($C = 24$ by default) of the UNet. Thus, the size of a voxel-wise convolution feature is $24 \times K$.

4) Point-Voxel Graph Neural Network: Given a point-voxel graph that takes both point-cloud features and high-dimension voxel features into account, we design a Point-Voxel Graph Neural Network (PV-GNN) to predict the category of each bronchial segment. The PV-GNN consists of Conv-Norm Blocks and a fully connected layer. A Conv-Norm Block is composed of a Mean Sage-Convolution (MSC) [52] layer, a Graph Normalization (GN) [53] layer, and a ReLU activation function. The MSC layer uses the mean aggregated function to merge information from node neighbors to overcome the inductive bias. Considering that adjacent nodes in the topology of the bronchial tree are relatively sparse, we build up a deep GNN for better information integrating point clouds. Since the GNN has a risk of suffering from gradient vanishing as it goes deeper, we introduce GN to shift and scale feature values, which makes graph neural networks converge much faster. Besides the first block, each block is followed by an element-wise addition that acts as a residual connection. Let H^k denote the output of the k-th block and σ denote the ReLU operation. The Conv-Norm block is formally defined as Eq. (2):

$$H^k = \sigma(\mathrm{GN}(\mathrm{MSC}(H^{k-1})) + H^{k-1}. \tag{2}$$

TABLE III

COMPARISON OF BRONCHUS CLASSIFICATION METHODS. LP [31], TS-NN [7], SGNET [44] ARE THE PREVIOUS STATE-OF-THE-ARTS METHODS. THE GNN-BASED METHODS (I.E., GCN [33], GAT [34], GIN [35]) UTILIZE **OUR PROPOSED POINT-VOXEL FEATURE** FOR GRAPH CONSTRUCTION. *p-Value* IS CALCULATED BETWEEN OUR MODEL AND OTHER MODELS UNDER THREE DIFFERENT SEEDS

Methods	Accuracy	Precision	Recall	F1-score	*p-value*	Flops (G)	Running Time (s)	Throughput (samples/s)
LP [31]	$81.8_{\pm 4.3}$	$74.7_{\pm 3.4}$	$79.2_{\pm 4.8}$	$77.0_{\pm 4.2}$	<0.001	-	0.75	1.33
TS-NN [51]	$76.8_{\pm 5.3}$	$77.8_{\pm 4.4}$	$74.9_{\pm 6.1}$	$76.2_{\pm 5.2}$	<0.001	0.38	0.04	24.99
SGNet [44]	$85.6_{\pm 5.3}$	$85.0_{\pm 5.2}$	$84.4_{\pm 6.2}$	$84.7_{\pm 5.6}$	<0.001	477.47	4.27	0.28
GCN [33]	$90.7_{\pm 5.7}$	$90.2_{\pm 4.6}$	$90.4_{\pm 5.6}$	$90.4_{\pm 5.2}$	0.006	13.21	2.52	0.40
GIN [35]	$90.0_{\pm 4.2}$	$89.2_{\pm 5.3}$	$90.1_{\pm 5.8}$	$89.4_{\pm 5.5}$	<0.001	13.12	2.51	0.40
GAT [34]	$91.2_{\pm 5.6}$	$91.6_{\pm 4.9}$	$90.7_{\pm 5.4}$	$91.1_{\pm 5.2}$	0.020	14.21	2.60	0.39
BCNet	$94.2_{\pm 3.8}$	$93.4_{\pm 6.3}$	$94.0_{\pm 4.9}$	$93.7_{\pm 5.5}$	-	13.07	1.00	1.00

5) CE Loss With Neighborhood Consistency Regularization:

Considering the topology prior that adjacent branches in a bronchial tree more likely belong to the same category, we design a novel Neighborhood Consistency Regularization (NCR) to penalize local spatial variations and force the nearby nodes of the same category to be closer in a latent space. Let $Y = \{y_1, y_2, \ldots, y_N\}$ be the set that contains the one-hot vector of the ground truth of each branch and $Z = \{z_1, z_2, \ldots, z_N\}$ be the set that contains the logit vector of the model prediction of each branch. The NCR loss is formulated as Eq. (3):

$$L_{NCR} = \frac{\sum_{i=1}^{N} \sum_{j}^{V_i} ||z_i - z_j|| \mathbb{I}(y_i = y_j)}{M}, \tag{3}$$

where V_i is the set of the i-th node's neighbors, j denotes the j-th node in this set, z_i denotes the logit vector of the i-th node from the output of the last fully connected layer, $\mathbb{I}(\cdot)$ is an indicator function that returns 1 when the condition is met and returns 0 otherwise, and M, N are the numbers of edges and nodes in the graph, respectively. Let α be a scalar to balance the weight of the regularization and CE loss, the overall loss function contains the above NCR and a vanilla Cross-Entropy loss and is formulated as Eq. (4):

$$L_{pvgnn} = L_{CE}(y_{pred}, y_{gt}) + \alpha L_{NCR}(y_{pred}), \tag{4}$$

where L_{pvgnn} has the same number of classes with L_{unet}.

D. Overall Loss for Bronchus Classification

We train our proposed model with the UNet and PV-GNN components in an end-to-end manner, since the gradients of the GNN branch can sent backward to the UNet branch through the operation of convolutional voxel feature extraction. Let β be a scalar to balance the weight of two tasks, the final loss L_{cls} is composed of two parts, which could be formulated as Eq. (5):

$$L_{cls} = L_{unet} + \beta L_{pvgnn}. \tag{5}$$

V. EXPERIMENTS AND RESULTS

A. Implementation Details

PyTorch 1.10 is used to build the model. All models are trained with an NVIDIA V100 GPU of 32GB. BronAtlas is split into the training/validation/testing set with 63/7/30 samples. We use the validation set to tune our parameters select the best-performed model and further test it on the test set. We augment the training data by applying random affine transformation and elastic deformation 99 times. Following the aforementioned data augmentation process, we transformed the samples into a cube with a shape of $128 \times 128 \times 128$. This standardized cube was used to train all the algorithms in this work. For the testing phase, we reverted the cube to its original shape for evaluation and visualization. By doing this, we were able to ensure a fair comparison across all the methods. We use DropEdge [54] for the classification model training to avoid over-fitting. The model is trained with the Adam optimizer, a batch size of 128, and a learning rate of 0.001 for 50 epochs. The number of layers and hidden dimensions in PV-GNN is set to 5 and 256, respectively.

We follow [31] to adopt four metrics: accuracy, precision, recall, and F1-score. The overall F1-score is computed by taking the average of the F1-scores of all classes, while the F1-score of each class is based on the harmonic mean of the precision and recall. All performance metrics reported herein are computed using the One-versus-All (OvA) strategy. OvA treats one class as the positive class and combines all other classes into a single negative class for each classification task, thereby enabling the evaluation of classifier performance on a per-class basis.

TABLE IV

FIVE-FOLD CROSS-VALIDATION RESULT OF THREE REPRESENTATIVE METHODS. IN THIS SETTING, WE USE 56 SAMPLES FOR TRAINING, 14 SAMPLES FOR VALIDATION, AND 30 SAMPLES FOR TESTING

Method	Accuracy	Precision	Recall	F1-score	p-value
CNN	$81.7_{\pm 0.3}$	$82.3_{\pm 0.2}$	$81.1_{\pm 0.4}$	$81.7_{\pm 0.4}$	<0.001
GIN	$87.6_{\pm 0.1}$	$87.1_{\pm 0.2}$	$86.3_{\pm 0.1}$	$86.7_{\pm 0.2}$	<0.001
BCNet	$93.2_{\pm 0.1}$	$93.0_{\pm 0.1}$	$92.6_{\pm 0.1}$	$92.8_{\pm 0.1}$	-

TABLE V

DETAILED PERFORMANCE ANALYSIS ON THE OUT-OF-DOMAIN DATA FROM ANOTHER DATA CENTER WHICH CONTAINS 30 CASES. *p-Value* IS CALCULATED BETWEEN OUR MODEL AND OTHER MODELS UNDER THREE DIFFERENT SEEDS

Method	Accuracy	Precision	Recall	F1-score	p-value
CNN	$84.5_{\pm 7.0}$	$83.1_{\pm 7.1}$	$86.8_{\pm 6.3}$	$84.9_{\pm 6.7}$	<0.001
GIN	$88.7_{\pm 0.6}$	$88.0_{\pm 5.7}$	$91.4_{\pm 5.4}$	$89.6_{\pm 5.4}$	<0.01
BCNet	$91.7_{\pm 4.6}$	$90.3_{\pm 4.9}$	$94.3_{\pm 3.5}$	$92.3_{\pm 4.1}$	-

B. Comparison With the State-of-the-Art Methods

The quantitative comparison for bronchial classification is shown in Table III. The previous state-of-the-art methods,

TABLE VI
PERFORMANCE ON THE ABNORMAL CATEGORY IN BRONCHUS ATLAS BENCHMARK. RESULTS ARE TESTED USING THE BEST-PERFORMING MODEL IN THE VALIDATION SET

Method	Accuracy	Precision	Recall	F1-score	p-value
CNN	24.5	26.7	15.4	19.5	<0.001
GIN	39.3	38.2	22.0	27.9	<0.001
BCNet	61.2	60.7	50.0	54.8	-

TABLE VII
ABLATION STUDY OF DIFFERENT COMPONENTS USED IN OUR BCNET. 'PF', 'VF', AND 'NCR' DENOTE POINT FEATURE, VOXEL FEATURE, AND NEIGHBORHOOD CONSISTENCY REGULARIZATION, RESPECTIVELY. NOTE THAT ALL METHODS INVOLVE A VOXEL FEATURE NEEDED TO TRAIN THE UNET. *p-Value* IS CALCULATED BETWEEN OUR MODEL AND OTHER MODELS UNDER THREE DIFFERENT SEEDS

Methods	CNN	GNN			F1-score	p-value
		PF	VF	NCR		
CNN	✓				$85.7_{\pm 8.5}$	<0.001
V-GNN			✓		$91.4_{\pm 5.0}$	0.04
P-GNN		✓			$88.2_{\pm 5.7}$	<0.001
VN-GNN			✓	✓	$91.7_{\pm 5.5}$	0.06
PN-GNN		✓		✓	$90.7_{\pm 5.9}$	0.01
PV-GNN		✓	✓		$91.8_{\pm 6.0}$	0.07
PVN-GNN		✓	✓	✓	$92.4_{\pm 6.5}$	0.16
BCNet (GNN)	✓	✓	✓	✓	$91.2_{\pm 6.3}$	0.03
BCNet (CNN)	✓	✓	✓	✓	$93.7_{\pm 5.5}$	-

we implement them by following the description in their paper and using the same training setting as our work. For the GNN-based methods, we first pre-train a UNet to segment the bronchus and extract the dimension voxel feature, then use the GNN that takes the point cloud feature and high-dimension voxel feature as the node feature for bronchus classification. Our final BCNet is trained using a multi-task learning pipeline with the PV-GNN and CNN. For inference, BCNet solely relies on the CNN part.

Observations from Table III indicate that our uniquely designed point-voxel feature allows the Graph Neural Networks (GNNs) to deliver superior results compared to previously established state-of-the-art methods. This shows that compared with the previous handcraft feature [31] or other graph construction feature [44], the point cloud feature and convolutional feature are quite important for the distinguishable graph representation learning. Our Bronchus Classification Network (BCNet) significantly outperforms the previously established state-of-the-art SGNet, achieving an improvement of 8.8% in accuracy and 9.0% in the F1-score. This underscores the substantial effectiveness of our proposed structure-guided learning framework. Notably, BCNet exhibits an advantage in efficiency over Graph Neural Network (GNN)-based methods. While the latter requires both Convolutional Neural Network (CNN) and GNN components during the inference stage, BCNet only needs the CNN part for inference, leading to greater efficiency and compromising performance. We present the outcomes of representative methods in Table IV, conducted under a five-fold cross-validation setting on our dataset. To achieve five-fold cross-validation, we adopt the data partition of "56/14/30". Note that, for the data partition of "63/7/30" in Section III-C, we do not perform cross-validation. Additionally, we further evaluate our model's performance on out-of-domain data, as shown in Table V. All these results support that our methods are effective.

We also provide the averaging result of the abnormal cases, and it is shown in Table VI. We can see that our method achieves the highest performance and significantly exceeds the previous methods on the abnormal samples by over 25% on the F1 scores. The abnormal cases are clinically important. However, due to the limited abnormal cases in the training set, there is still a gap to achieve better performance, which will be our future work.

C. Ablation Study

The ablation study for the classification task is in Table VII. 'CNN' uses only a UNet to label the input binary mask, which shows inferior results due to the lack of structure modeling. '*-GNN' denotes the method that adopts a graph neural network with different node representations. 'PF' denotes the point-wise coordinate feature based on point clouds, while 'VF' means the voxel-wise convolution feature. 'NCR' denotes Neighborhood Consistency Regularization. Using the graph representation of point clouds, 'P-GNN' outperforms the 'CNN' by 6.3% on the F1-score. 'PV-GNN' means the proposed point-voxel graph neural network that adopts both the voxel-wise convolutional features and the point cloud features as the node representations, and it can capture the texture and diameter information, thus significantly exceeding the 'P-GNN' by 3.8% F1-score. 'PN-GNN' denotes the GNN that employs the point cloud features as node features and utilizes the NCR. 'PVN-GNN' means integrating the NCR with the 'PV-GNN' model, which brings a performance gain of 0.6%.

'BCNet (GNN)' and 'BCNet (CNN)' denote the results of the CNN (UNet) and GNN branches in our BCNet, respectively. Under the guidance of the GNN, the CNN branch in BCNet obtains 93.7% F1-score and impressively outperforms the CNN baseline (85.7%) by **8% on F1-score**. The improvement well verifies the idea of structure-guided learning that training the GNN branch can boost the feature learning and classification result of the CNN branch. Of the reasons the performance of the GNN part degrades, one is that the GNN result needs to be recovered to the cube with an unlearnable region growth algorithm. The reason for PVN-GNN is better than BCNet (GNN) on F1-score is that PVN-GNN uses a pre-trained and fixed classification CNN network to extract static voxel features, which are used to initialize the nodes in PVN-GNN. In this situation, the voxel features are stable and discriminative for classification. In contrast, BCNet (GNN) uses voxel features extracted from a CNN which is updated in real-time, which may introduce unstable noises and negatively affect the initial training process of the GNN branch, leading to degraded performance. Note that 'BCNet (GNN)' still achieves the state-of-the-art result among existing methods in Table III. Interestingly, during the inference, BCNet can run only the CNN branch without relying on the GNN branch and

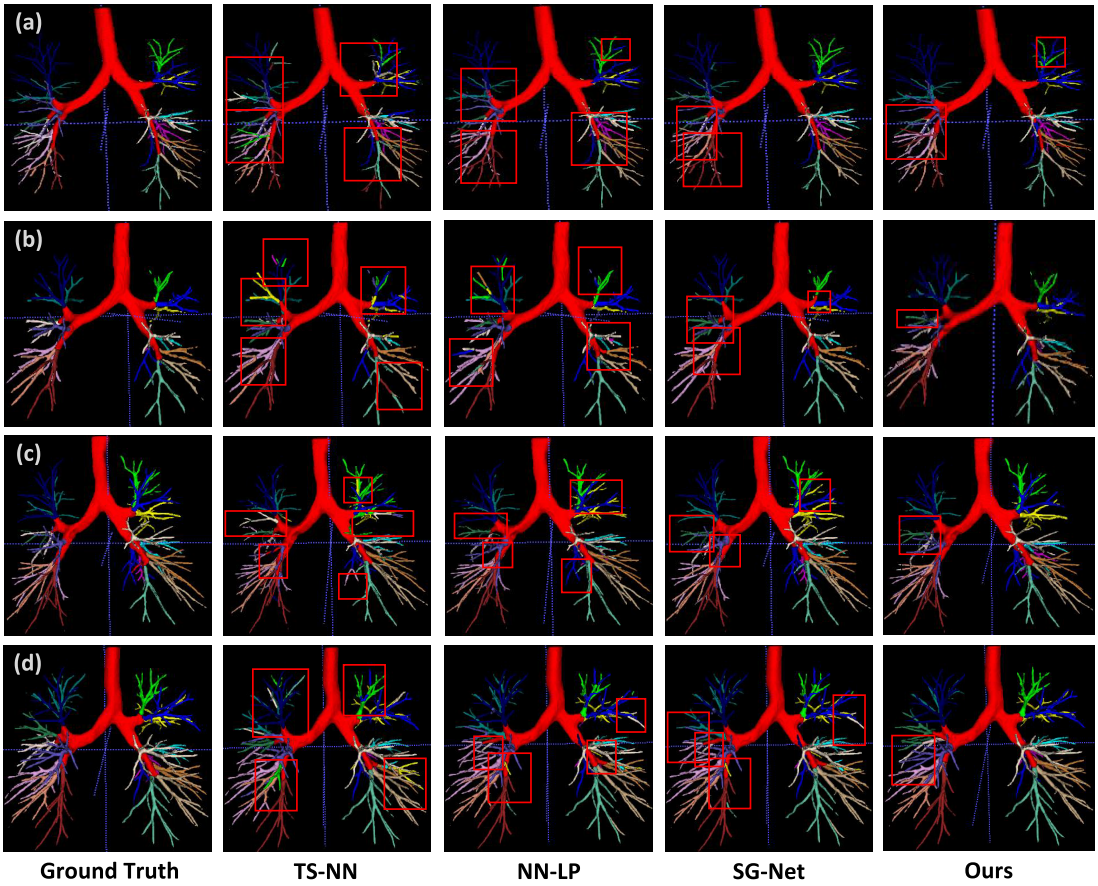

Fig. 5. Qualitative analysis on the BronAtlas benchmark. The misclassified bronchus is bounded by red boxes. TS-NN [7] suffers from errors related to branching variability. LP [31] fails to distinguish the segments with similar angles. SGNet [44] misclassifies the thin bronchus while the proposed BCNet shows robust predictions.

thus enjoys both the best performance 93.7%, and the same efficiency as the CNN baseline.

VI. Discussion

A. Constrains and Limitations

The one-stage framework utilizes UNet for direct multi-class segmentation from CT images and processes samples in approximately 3.5 seconds each. Despite its speed, the performance metrics show significant limitations, with an accuracy of only 36.1 ($\pm$15.7), precision of 34.3 ($\pm$12.2), and particularly low recall of 11.8 ($\pm$3.7), resulting in an F1-score of 17.0 ($\pm$4.9), as shown in TABLE VIII. The p-value < 0.001 highlights the considerable performance gap compared to the two-stage approach. Our two-stage framework, though more time-consuming, clearly outperforms the one-stage. In the training process, we start with UNet generating a binary mask in 3.5 seconds per sample, followed by construction of bronchial tree (1.5 seconds per sample) and CNN-GNN based classification (1.01 seconds per sample), totaling around 6 seconds per sample. This approach attains an accuracy of 93.6 ($\pm$5.6), precision of 93.4 ($\pm$6.1), and recall of 91.7 ($\pm$8.2), with an F1-score of 92.5 ($\pm$7.0). It is worth noting that during the inference stage, we only need 1 second to obtain the classification result from the CNN branch without prepossessing like building a bronchial tree.

TABLE VIII
Comparison of One-Stage and Two-Stage Frameworks for CT Image Segmentation. The One-Stage Approach Employs UNet for Direct Multi-Class Segmentation From CT Images. The Two-Stage Framework First Utilizes UNet to Generate a Binary Mask From the CT, Followed by the Application of BCNet for Multi-Class Classification

Method	Accuracy	Precision	Recall	F1-score	p-value
one-stage	$36.1_{\pm 15.7}$	$34.3_{\pm 12.2}$	$11.8_{\pm 3.7}$	$17.0_{\pm 4.9}$	<0.001
two-stage	$93.6_{\pm 5.6}$	$93.4_{\pm 6.1}$	$91.7_{\pm 8.2}$	$92.5_{\pm 7.0}$	-

The one-stage method's failure is largely due to its inability to separately optimize segmentation and classification, forcing UNet to manage both tasks simultaneously. This can lead to conflicts between segmentation accuracy and classification precision. In contrast, the two-stage method allows each network (UNet for segmentation and BCNet for classification) to specialize and optimize its respective task, resulting in significantly enhanced overall performance. In summary, the two-stage framework, despite its longer processing time, offers a substantial increase in performance, making it preferable for applications where accuracy is critical. The one-stage method, although faster, suffers from significant performance deficits, rendering it less suitable for high-stakes clinical applications.

TABLE IX
ANALYSIS ON THE HYPER-PARAMETERS α AND β. PARAMETER α IS USED IN THE NEIGHBORHOOD CONSISTENCY REGULARIZATION (NCR) TO BALANCE THE TRADE-OFF BETWEEN CROSS-ENTROPY LOSS AND NCR, AS SHOWN IN EQ. 4. PARAMETER β ADJUSTS THE TRADE-OFF BETWEEN THE CROSS-ENTROPY LOSS ON THE CNN AND GNN BRANCHES, AS DETAILED IN EQ. 5

α	0	0.5	1	1.5	2
F1-score	$92.8_{\pm 6.1}$	$93.1_{\pm 5.8}$	$93.7_{\pm 5.5}$	$93.3_{\pm 4.9}$	$93.0_{\pm 6.5}$
β	-	0.5	1	2	-
F1-score	-	$93.5_{\pm 4.7}$	$93.7_{\pm 5.5}$	$93.4_{\pm 5.3}$	-

TABLE X
ANALYSIS ON THE NUMBER OF POINTS K IN THE GRAPH CONSTRUCTION PROCESS

K	6	8	10	12	14
F1-score	$93.3_{\pm 5.8}$	$93.4_{\pm 5.9}$	$93.7_{\pm 5.5}$	$93.2_{\pm 5.2}$	$93.5_{\pm 6.1}$

Another obstacle in our current work is the performance of the abnormal category predictions. Despite BCNet being the best-performing model with an accuracy of 61.2 and F1-score of 54.8, it, along with CNN and GIN, shows substantial room for improvement, particularly in recall rates. Future efforts will focus on refining these models through advanced neural architectures, increased data augmentation, sophisticated feature engineering, and hybrid model approaches. Moreover, implementing advanced training techniques such as transfer learning could also enhance our model's ability to accurately identify abnormal categories, addressing the critical need for reliable detection in medical imaging. These strategies are pivotal as we aim to close the performance gap between abnormal and normal category predictions.

B. Sensitive and Qualitative Analysis

The sensitivity analysis for our BCNet architecture, detailed in Table IX, evaluates the impact of hyper-parameters α and β. The parameter α, which governs the balance between cross-entropy loss and neighborhood consistency regularization (NCR) as per Eq. 4, shows that increasing α from 0 to 1 enhances the F1-score, reaching a peak at $\alpha = 1$ (93.7). This indicates optimal NCR application, improving model generalization. Beyond this, further increases in α slightly degrade performance, suggesting over-regularization. Similarly, the parameter β, adjusting the cross-entropy loss balance between the CNN and GNN branches as outlined in Eq. 5, reveals that both $\beta = 0.5$ and $\beta = 1$ optimize the model's performance with F1-scores of 93.5 and 93.7, respectively. A shift to $\beta = 2$ results in a minor decrease in the F1-score to 93.4, indicating less effective loss distribution.

Table X demonstrates the effect of varying K on the F1 score. As K increases from 6 to 10, the F1 score improves, reaching a peak of 93.7 at $K = 10$. However, further increases in K result in marginal changes, suggesting an optimal K range of 1 to 1.5 for achieving the highest F1 score. We also provide the qualitative comparison result for bronchus segmentation in Fig. 5. As we can see, our method produces less misclassified bronchus (marked in red box) than other approaches.

VII. CONCLUSION

In this paper, we present the BCNet, a structure-guided learning framework for bronchus classification based on CT imaging data. With the simultaneous training of our proposed Point-Voxel GNN at the segment level, the BCNet learns to harvest structure-aware representations for the shared convolutional backbone and hence yields better predictions at the voxel-level branch. Besides, we exploit the structure prior with a novel neighborhood consistency regularization to boost the performance. We contribute the BronAtlas benchmark that contains 100 CT scans with voxel-wise masks and segment-level labels to facilitate future research. The experimental results on the BronAtlas benchmark show that our proposed model significantly exceeds the state-of-the-art methods.

REFERENCES

[1] J. P. Sieren et al., "SPIROMICS protocol for multicenter quantitative computed tomography to phenotype the lungs," *Amer. J. Respiratory Crit. Care Med.*, vol. 194, no. 7, pp. 794–806, Oct. 2016.
[2] Y. Qin et al., "AirwayNet: A voxel-connectivity aware approach for accurate airway segmentation using convolutional neural networks," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2019, pp. 212–220.
[3] Y. Kawata, "Representation of thoracic n1 lymph nodes group in contrast-enhanced CT images using distance maps based on tracheobronchial labeling," in *Medical Imaging 2023: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 12468. Bellingham, WA, USA: SPIE, 2023, pp. 368–373.
[4] H. Li, Z. Tang, Y. Nan, and G. Yang, "Human treelike tubular structure segmentation: A comprehensive review and future perspectives," *Comput. Biol. Med.*, vol. 151, Dec. 2022, Art. no. 106241.
[5] J. Yun et al., "Improvement of fully automated airway segmentation on volumetric computed tomographic images using a 2.5 dimensional convolutional neural net," *Med. Image Anal.*, vol. 51, pp. 13–20, Jan. 2019.
[6] Q. Meng, H. R. Roth, T. Kitasaka, M. Oda, J. Ueno, and K. Mori, "Tracking and segmentation of the airways in chest ct using a fully convolutional network," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2017, pp. 198–207.
[7] S. A. Nadeem, "A fully automated CT-based airway segmentation branch labeling algorithm using deep learn conventional image process," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Iowa, Iowa City, IA, USA, 2020.
[8] M. Wang et al., "Automated labeling of the airway tree in terms of lobes based on deep learning of bifurcation point detection," *Med. Biol. Eng. Comput.*, vol. 58, no. 9, pp. 2009–2024, Sep. 2020.
[9] K. Mori, J. Hasegawa, Y. Suenaga, and J. Toriwaki, "Automated anatomical labeling of the bronchial branch and its application to the virtual bronchoscopy system," *IEEE Trans. Med. Imag.*, vol. 19, no. 2, pp. 103–114, Jul. 2000.
[10] M. Sato, T. Murayama, and J. Nakajima, "Concepts and techniques: How to determine and identify the appropriate target segment in anatomical pulmonary segmentectomy?" *J. Thoracic Disease*, vol. 11, no. 3, pp. 972–986, Mar. 2019.
[11] R. Keuth, M. Heinrich, M. Eichenlaub, and M. Himstedt, "Airway label prediction in video bronchoscopy: Capturing temporal dependencies utilizing anatomical knowledge," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 19, no. 4, pp. 713–721, Jan. 2024.
[12] H. Gong et al., "Multi-task learning for thyroid nodule segmentation with thyroid region prior," in *Proc. IEEE 18th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2021, pp. 257–261.
[13] H. Gong, J. Chen, G. Chen, H. Li, G. Li, and F. Chen, "Thyroid region prior guided attention for ultrasound segmentation of thyroid nodules," *Comput. Biol. Med.*, vol. 155, Mar. 2023, Art. no. 106389.
[14] M. Zhang et al., "Multi-site, multi-domain airway tree modeling," *Med. Image Anal.*, vol. 90, Jul. 2023, Art. no. 102957.
[15] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351. Cham, Switzerland: Springer, 2015, pp. 234–241.

[16] A. Heitz et al., "Lubrav: A new framework for the segmentation of the Lung's tubular structures," in *Proc. IEEE 18th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2021, pp. 948–952.
[17] J.-P. Charbonnier, E. M. V. Rikxoort, A. A. A. Setio, C. M. Schaefer-Prokop, B. V. Ginneken, and F. Ciompi, "Improving airway segmentation in computed tomography using leak detection with convolutional networks," *Med. Image Anal.*, vol. 36, pp. 52–60, Feb. 2017.
[18] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "NnU-net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021.
[19] Y. Qin et al., "Learning bronchiole-sensitive airway segmentation CNNs by feature recalibration and attention distillation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*, 2020, pp. 221–231.
[20] Y. Qin et al., "Learning tubule-sensitive CNNs for pulmonary airway and artery-vein segmentation in CT," *IEEE Trans. Med. Imag.*, vol. 40, no. 6, pp. 1603–1617, Jun. 2021.
[21] A. Garcia-Uceda, R. Selvan, Z. Saghir, H. A. W. M. Tiddens, and M. de Bruijne, "Automatic airway segmentation from computed tomography using robust and efficient 3-D convolutional neural networks," *Sci. Rep.*, vol. 11, no. 1, pp. 1–15, Aug. 2021.
[22] H. Zheng et al., "Alleviating class-wise gradient imbalance for pulmonary airway segmentation," *IEEE Trans. Med. Imag.*, vol. 40, no. 9, pp. 2452–2462, Sep. 2021.
[23] H. Gong et al., "Intensity confusion matters: An intensity-distance guided loss for bronchus segmentation," in *Proc. IEEE Int. Conf. Multimedia Expo. (ICME)*, Jun. 2024, pp. 1–22.
[24] H. Zheng et al., "Refined local-imbalance-based weight for airway segmentation in CT," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*, 2021, pp. 410–419.
[25] Z. Tang, Y. Nan, S. Walsh, and G. Yang, "Adversarial transformer for repairing human airway segmentation," *IEEE J. Biomed. Health Informat.*, vol. 1, no. 1, pp. 1–9, Apr. 2023.
[26] Y. Nan et al., "Fuzzy attention neural network to tackle discontinuity in airway segmentation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 1, no. 1, pp. 1–14, Apr. 2023.
[27] J. Tschirren, G. McLennan, K. Palagyi, E. A. Hoffman, and M. Sonka, "Matching and anatomical labeling of human airway tree," *IEEE Trans. Med. Imag.*, vol. 24, no. 12, pp. 1540–1547, Dec. 2005.
[28] A. Feragen et al., "A hierarchical scheme for geodesic anatomical labeling of airway trees," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI*. Cham, Switzerland: Springer, 2012, pp. 147–155.
[29] Y. Liu et al., "Automated anatomical labeling of a topologically variant abdominal arterial system via probabilistic hypergraph matching," *Med. Image Anal.*, vol. 75, Jan. 2022, Art. no. 102249.
[30] P. Lo and E. M. van Rikxoort, "A bottom-up approach for labeling of human airway trees," in *Proc. MICCAI Int. WS. Pulm. IM. Anal*, 2011, pp. 1–20.
[31] T. Zhao, Z. Yin, J. Wang, D. Gao, Y. Chen, and Y. Mao, "Bronchus segmentation and classification by neural networks and linear programming," in *Proc. Int. Conf. Med. Image Comput. Computer-Assisted Intervent*, 2019, pp. 230–239.
[32] W. Yu et al., "TNN: Tree neural network for airway anatomical labeling," *IEEE Trans. Med. Imag.*, vol. 42, no. 1, pp. 103–118, Jan. 2023.
[33] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Proc. NeurIPS*, 2016, pp. 1–22.
[34] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," in *Proc. Int. Conf. Learn. Represent.*, 2018, pp. 1–15.
[35] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?" in *Proc. ICLR*, 2019, pp. 1–19.
[36] X. Song, J. Li, and X. Qian, "Diagnosis of glioblastoma multiforme progression via interpretable structure-constrained graph neural networks," *IEEE Trans. Med. Imag.*, vol. 42, no. 2, pp. 380–390, Feb. 2023.
[37] Y. Li et al., "A hierarchical graph V-Net with semi-supervised pretraining for histological image based breast cancer classification," *IEEE Trans. Med. Imag.*, vol. 42, no. 12, pp. 3907–3918, Dec. 2023.
[38] B. Liang, H. Gong, L. Lu, and J. Xu, "Risk stratification and pathway analysis based on graph neural network and interpretable algorithm," *BMC Bioinf.*, vol. 23, no. 1, p. 394, Sep. 2022.
[39] H. Gong et al., "Unbiased curriculum learning enhanced global-local graph neural network for protein thermodynamic stability prediction," *Bioinformatics*, vol. 39, no. 10, p. 589, Oct. 2023.
[40] G. Corso, H. Stark, S. Jegelka, T. Jaakkola, and R. Barzilay, "Graph neural networks," *Nature Rev. Methods Primers*, vol. 4, no. 1, p. 17, 2024.
[41] A. G.-U. Juarez, R. Selvan, Z. Saghir, and M. de Bruijne, "A joint 3D UNet-graph neural network-based method for airway segmentation from chest CTs," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2019, pp. 583–591.
[42] R. Selvan et al., "Graph refinement based airway extraction using mean-field networks and graph neural networks," *Med. Image Anal.*, vol. 64, Aug. 2020, Art. no. 101751.
[43] T. Zhao and Z. Yin, "Airway anomaly detection by prototype-based graph neural network," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*, 2021, pp. 195–204.
[44] Z. Tan, J. Feng, and J. Zhou, "SGNet: Structure-aware graph-based network for airway semantic segmentation," in *Proc. Int. Conf. Med. Image Comput. Computer-Assisted Intervent*, 2021, pp. 153–163.
[45] W. Lou et al., "Structure embedded nucleus classification for histopathology images," *IEEE Trans. Med. Imag.*, vol. 1, no. 1, pp. 1–16, Jun. 2024.
[46] W. Lou, G. Li, X. Wan, and H. Li, "Cell graph transformer for nuclei classification," in *Proc. 38th AAAI Conf. Artif. Intell.*, 2024, pp. 3873–3881.
[47] P. Lo et al., "Extraction of airways from CT (EXACT'09)," *IEEE Trans. Med. Imag.*, vol. 31, no. 11, pp. 2093–2107, Nov. 2012.
[48] S. G. Armato et al., "The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans," *Med. Phys.*, vol. 38, no. 2, pp. 915–931, Feb. 2011.
[49] J. Williams, "The declaration of Helsinki and public health," *Bull. World Health Org.*, vol. 86, no. 8, pp. 650–651, Aug. 2008.
[50] T. C. Lee, R. L. Kashyap, and C. N. Chu, "Building skeleton models via 3-D medial surface axis thinning algorithms," *CVGIP, Graph. Models Image Process.*, vol. 56, no. 6, pp. 462–478, Nov. 1994.
[51] S. A. Nadeem et al., "A CT-based automated algorithm for airway segmentation using freeze-and-grow propagation and deep learning," *IEEE Trans. Med. Imag.*, vol. 40, no. 1, pp. 405–418, Jan. 2021.
[52] W. L. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–10.
[53] T. Cai, S. Luo, K. Xu, D. He, T.-Y. Liu, and L. Wang, "GraphNorm: A principled approach to accelerating graph neural network training," in *Proc. ICML*, 2021, pp. 1–29.
[54] Y. Rong, W. B. Huang, T. Xu, and J. Huang, "DropEdge: Towards deep graph convolutional networks on node classification," in *Proc. ICLR*, 2020, pp. 1–17.