

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*, Hong Shen

Abstract—CT-based bronchial tree analysis is a crucial step for the diagnosis of lung and airway diseases. However, the varying topology of bronchial trees across individuals presents a significant challenge to automatic bronchus classification. To address this issue, we propose the Bronchus Classification Network (BCNet), a structure-guided framework that leverages segment-level topological information using point clouds to learn voxel-level features. BCNet consists of two branches: a Point-Voxel Graph Neural Network (PV-GNN) for segment classification and a Convolutional Neural Network (CNN) for voxel labeling. These two branches are trained simultaneously to learn topology-aware features for their shared backbone, while also allowing the CNN branch to be run independently during inference. This ensures that BCNet retains the same inference efficiency as its CNN baseline. Experimental results demonstrate that BCNet significantly surpasses the state-of-the-art methods by over 8.0% in F1-score for bronchus classification. Additionally, we introduce BronAtlas, an open-access benchmark for bronchus imaging analysis, featuring high-quality voxel-wise annotations of both anatomical and abnormal bronchial segments. The BronAtlas dataset provides a valuable resource for the research community, enabling the development and evaluation of advanced bronchial tree analysis methods for disease diagnosis and surgical planning. The benchmark dataset is publicly available at [link](#)¹.

Index Terms—Bronchus classification, Benchmark, CT imaging, Guidance-based distillation, Graph neural network, Convolution neural network Multi-task learning Point cloud.

Haifan Gong and Wenhao Huang contribute equally to this work. Guanbin Li and Haofeng Li are the corresponding authors. This work was supported in part by the Shenzhen Science and Technology Program JCYJ20220818103001002, in part by the Guangdong Provincial Key Laboratory of Big Data Computing, The Chinese University of Hong Kong, Shenzhen, in part by the Longgang District Special Funds for Science and Technology Innovation (LGKCSPT2023002), in part by the National Natural Science Foundation of China (No.62102267), and in part by the Guangdong Basic and Applied Basic Research Foundation (2023A1515011464).

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

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

¹<https://osf.io/pskr9/?viewonly=94fa3d87274b4095ac9a4b88cc9a1341>.

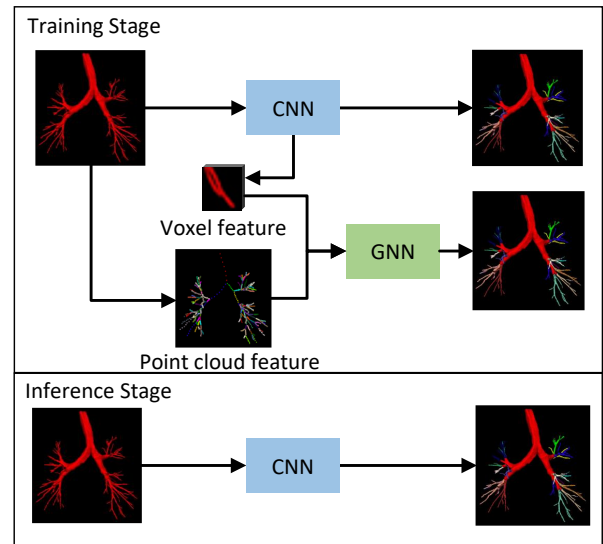


Fig. 1. The key idea of our BCNet. The upper part shows the training process involves the GNN and CNN. The GNN and CNN parts share the same voxel-wise CNN feature map, which is used to represent the bronchus in the GNN part. By computing loss in the GNN part, we encourage the shared feature map to better 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.

I. INTRODUCTION

Lung airway analysis based on CT imaging is clinically critical as it provides quantitative information that assists in lung disease diagnosis and surgical navigation [1]–[3]. The reconstruction of the bronchial tree, which serves as the foundation for lung airway analysis, typically involves two key steps [4]–[6]. The first step involves extracting a binary airway-tree mask from CT images, which delineates the structure of the bronchial tree. The second step is the classification of bronchus segments, where major anatomical branches are labeled accordingly. This automatic reconstruction of the bronchial tree supports various clinical processes, including the matching of individual airway tree phenotypes and the classification of lung lobes or segments [7], [8]. Furthermore, accurate labeling of bronchial segments is crucial

for clinicians as it helps identify which anatomical segments are affected by disease, thereby aiding in surgical planning and intervention [9]–[12]. Despite these advancements, automatic bronchus segment classification still faces a significant challenge: efficiently modeling and exploiting the complex topology of the bronchial tree. Addressing this challenge is essential for improving the accuracy and reliability of bronchial tree reconstruction, ultimately enhancing clinical outcomes.

Motivated by the significance of structural priors in medical image analysis [13], [14], we propose a novel framework, the Bronchus Classification Network (BCNet), designed to enhance the classification of bronchus segmentation from CT images. BCNet capitalizes on the intrinsic topology of the bronchial tree, incorporating a Point-Voxel Graph Neural Network (PV-GNN) to augment the representation learning capabilities of a Convolutional Neural Network (CNN). The core innovation of BCNet lies in its ability to employ segment-wise topological information through the GNN to inform and guide the learning of voxel-wise convolutional features via the CNN. This synergistic approach enables the CNN component to more accurately distinguish bronchus segments at the voxel level, as depicted in Figure 1. Significantly, during inference, BCNet operates solely with the convolutional branch, obviating the need for the graph model. This allows BCNet to achieve the accuracy enhancements afforded by GNN training while preserving the computational efficiency characteristic of a CNN-based approach.

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

In recent years, deep learning (DL) has significantly advanced bronchus segmentation [4], [15]. Various single-stage networks, such as U-Net [16], 2.5D Net [5], [17], and 3D U-Net [2], [6], [12], [18]–[23], have been employed for this task. Meng *et al.* [6] proposed an image tracking method to enhance segmentation accuracy. Qin *et al.* [20], [21] developed CNN-based methods that demonstrate superior sensitivity to tenuous peripheral bronchioles. Zheng *et al.* [23] applied gradient erosion and dilation operators to mitigate inter-class imbalance issues in bronchus segmentation, while their subsequent work [24] introduced gradient ratio adjustment and weight enhancement strategies to handle hard-to-segment regions. Tang *et al.* [25] introduced an adversarial transformer network for improved airway segmentation, and Nan *et al.* [26] developed a fuzzy attention neural network aimed at addressing segmentation discontinuities. Additionally, Gong *et al.* [12] proposed an intensity-clarified loss function to better segment the bronchus from CT scans.

Bronchus classification [4], [27], [28] remains a challenging task due to the substantial variability in bronchial tree topology across individuals. This classification typically builds on the results of bronchus segmentation methods [2], [17]–[25]. Traditional methods for bronchus classification include geodesic matching [28] and probabilistic hypergraph matching [29]. Wang *et al.* [8] achieved lobar-level bronchus classification using a key-point detection approach, while Lo *et al.* [30] proposed a bottom-up rule-based method for bronchial labeling. In the realm of deep learning-based methods, Zhao *et al.* [31] conducted an analysis of the average inclination angle of bronchial segments in the training set and utilized linear programming to post-process the airway structures predicted by neural models. Nadeem *et al.* [7] introduced a two-stage neural network that labels the lobar-level and segment-level bronchus respectively, ensuring detailed and hierarchical classification. Further, Yu *et al.* [32] proposed the tree neural network for comprehensive airway labeling, demonstrating the efficacy of structured neural networks in capturing complex bronchial topologies.

B. Graph Neural Networks for Segmentation

Recently, Graph Neural Networks (GNNs) [33]–[35] have become increasingly important in the biomedical domain [36]–[40], particularly for addressing segmentation tasks [41]–[46]. For instance, Garcia-Uceda *et al.* [41] enhanced airway binary segmentation by replacing the deepest convolutional layer in a U-Net [16] 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 [16], 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

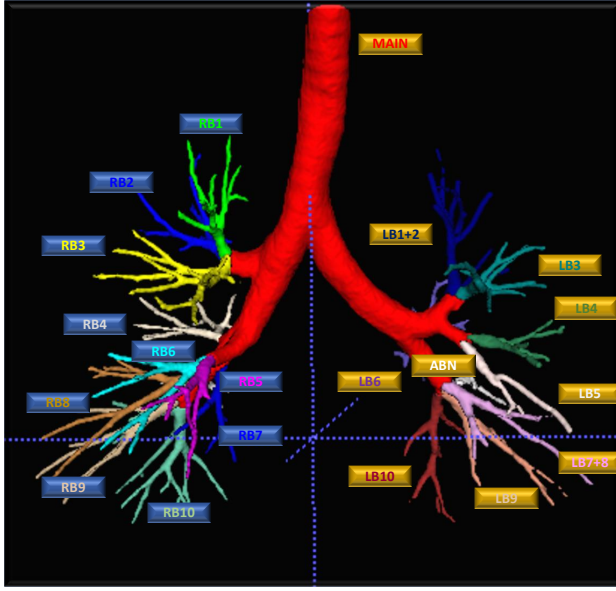


Fig. 2. The visualization results of the bronchus segment labels. The segmental bronchi from the right lung (i.e., RB1 to RB10) and segmental bronchi from the left lung (i.e., LB1+2, LB3, ..., LB7+8, LB9, LB10) are defined according to [27]. “MAIN” denotes the main trachea while “ABN” denotes the abnormal 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.

abnormal bronchus structures, showcasing the application of GNNs in identifying pathological changes.

However, these prior methods often rely heavily on hand-crafted features [7], [8], [31] or require additional labels [41], [44], which can limit their adaptability and efficiency. Moreover, they frequently struggle to manage the individual variability in bronchial structures effectively, which is crucial for accurate classification and segmentation. To address these limitations, there is a pressing need for an efficient framework that leverages both voxel-wise features and point-cloud topology for bronchus classification. Our proposed Bronchus Classification Network (BCNet) integrates the structural topology of the bronchial tree using a Point-Voxel Graph Neural Network (PV-GNN) to enhance the representation learning capabilities of a Convolutional Neural Network (CNN), thereby improving voxel-level feature discrimination and bronchus classification performance while maintaining computational efficiency.

III. BRONATLAS: A BENCHMARK FOR BRONCHUS ANALYSIS

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

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

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 collected 60 cases from the publicly available databases EXACT’09 [47] and LIDC [48], and an additional 40 cases from our cooperative hospital. For both the public samples and the newly collected ones, each CT scan was annotated by two experts through a two-step annotation process. Initially, the experts created binary masks for airway segmentation. Subsequently, they labeled 20 segmental bronchi at the voxel level based on the anatomical structure, as illustrated in Fig. 2.

Following the binary segmentation of the bronchus, the radiologists labeled 18 bronchial segments and the main trachea according to [27]. Specifically, 10 segmental bronchi were labeled in the right lung (RB1 to RB10), and 8 in the left lung (LB1+2, LB3, ..., LB7+8, LB9, LB10). During the annotation process, the radiologists ensured that, except for the abnormal branches, all bronchial segments were connected to the main bronchial trunk and that there were no issues of misclassification, under-classification, non-abnormal co-drying, or cross-color. Bronchial segments not fitting these categories were classified as abnormal bronchi. The process of data acquisition and annotation adhered to the principles outlined in the Declaration of Helsinki [49] and received approval from the institutional ethical committee.

C. Dataset Summary

For the development of automatic bronchus diagnosis, we introduce a new benchmark, BronAtlas, specifically designed for bronchus analysis. BronAtlas comprises 100 cases of lung CT imaging, each with voxel-level annotations for 20 segmental bronchi categories (10 from the right lung, 8 from the left lung, 1 for the main trachea, and 1 for abnormal

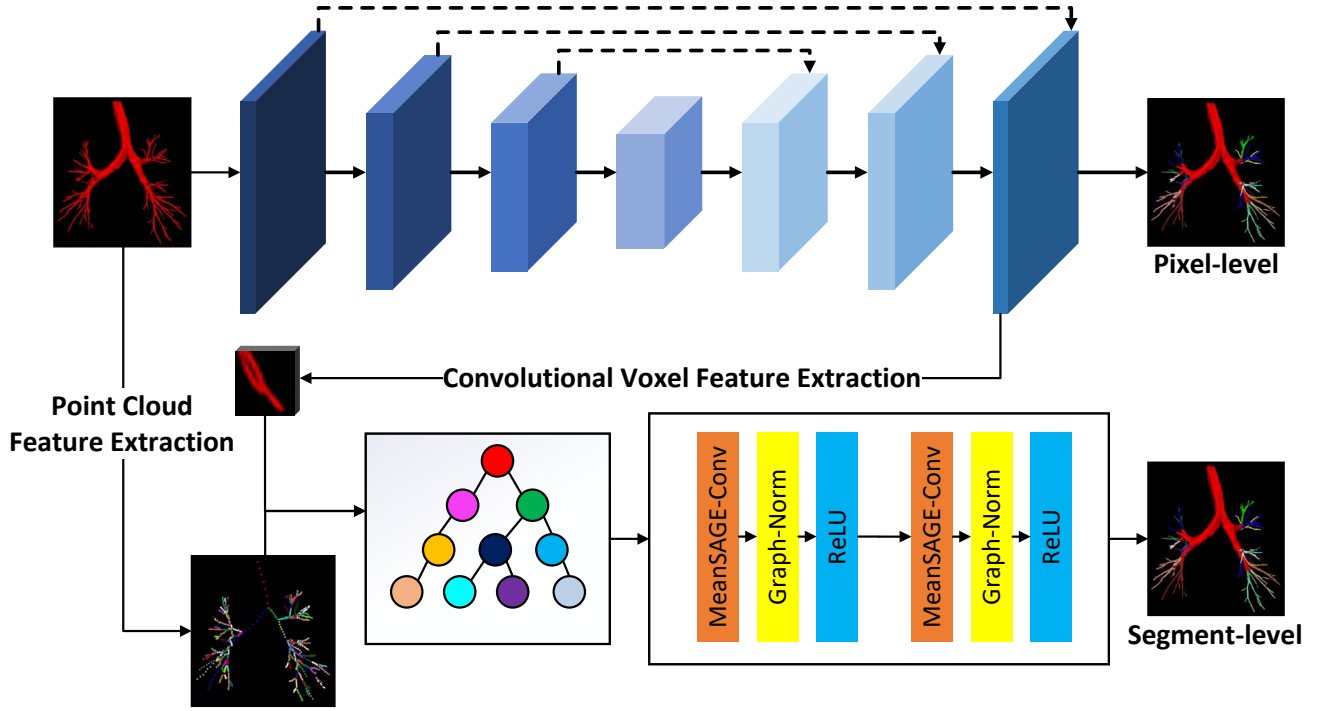


Fig. 3. Overview of the Bronchus Classification Network (BCNet). The upper part depicts a UNet that predicts voxel-level categorical labels of the bronchus based on a binary bronchus mask, refining segmentation through multiple convolutional layers. The lower part illustrates a Graph Neural Network (GNN) that operates at the segment level, predicting bronchus categories by leveraging structural and topological information. During training, gradients from the GNN are propagated backward through the UNet, guiding its representation learning and enhancing overall classification accuracy.

bronchus). The “abnormal bronchus segment” refers to congenital abnormal bronchus. The dataset includes 33 abnormal cases in the training and validation set, and 12 abnormal cases in the test set. A visualization of the 20 bronchial categories is provided in Figure 2, and a detailed summary of the samples in the benchmark is presented in Table I. The data from different sources were combined and split into training, validation, and test sets with 63, 7, and 30 cases, respectively.

The patient data used in this study was anonymized, with all personally identifiable information removed prior to model training and evaluation to ensure privacy. To minimize bias, we utilized a diverse dataset that includes CT scans from individuals of various ages, genders, and health conditions. Despite these efforts, we acknowledge the need for a more systematic approach to addressing bias, such as stratified sampling or bias audits, which we aim to implement in future work. This comprehensive and meticulously curated dataset aims to facilitate advancements in the automatic diagnosis of bronchial conditions.

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 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

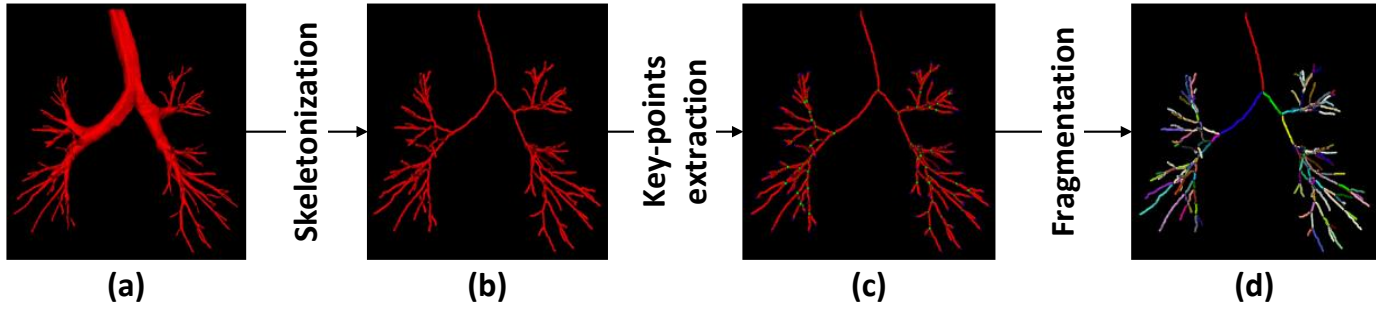


Fig. 4. Construction of bronchial tree: (1) skeletonize the bronchus; (2) extract the key points separating the bronchial fragments; (3) separate the bronchus into segments.

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 [19] 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:

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

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

a) 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.

b) 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 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.

c) 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 $24 \times K$.

d) 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

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

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	<i>p-value</i>	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

ReLU operation. The Conv-Norm block is formally defined as Eq. (2):

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

e) *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 $Y = \{y_1, y_2, \dots, y_N\}$ be the set that contains the one-hot vector of the ground truth of each branch and $Z = \{z_1, z_2, \dots, z_N\}$ be the set that contains the logit vector of the model prediction of each branch. The NCR loss is formulated as Eq. (3):

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

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):

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

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):

$$L_{cls} = L_{unet} + \beta L_{pvgnn}. \quad (5)$$

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

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	-

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.

B. Evaluation Metrics

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.

C. 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, 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.

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	-

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.

D. 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

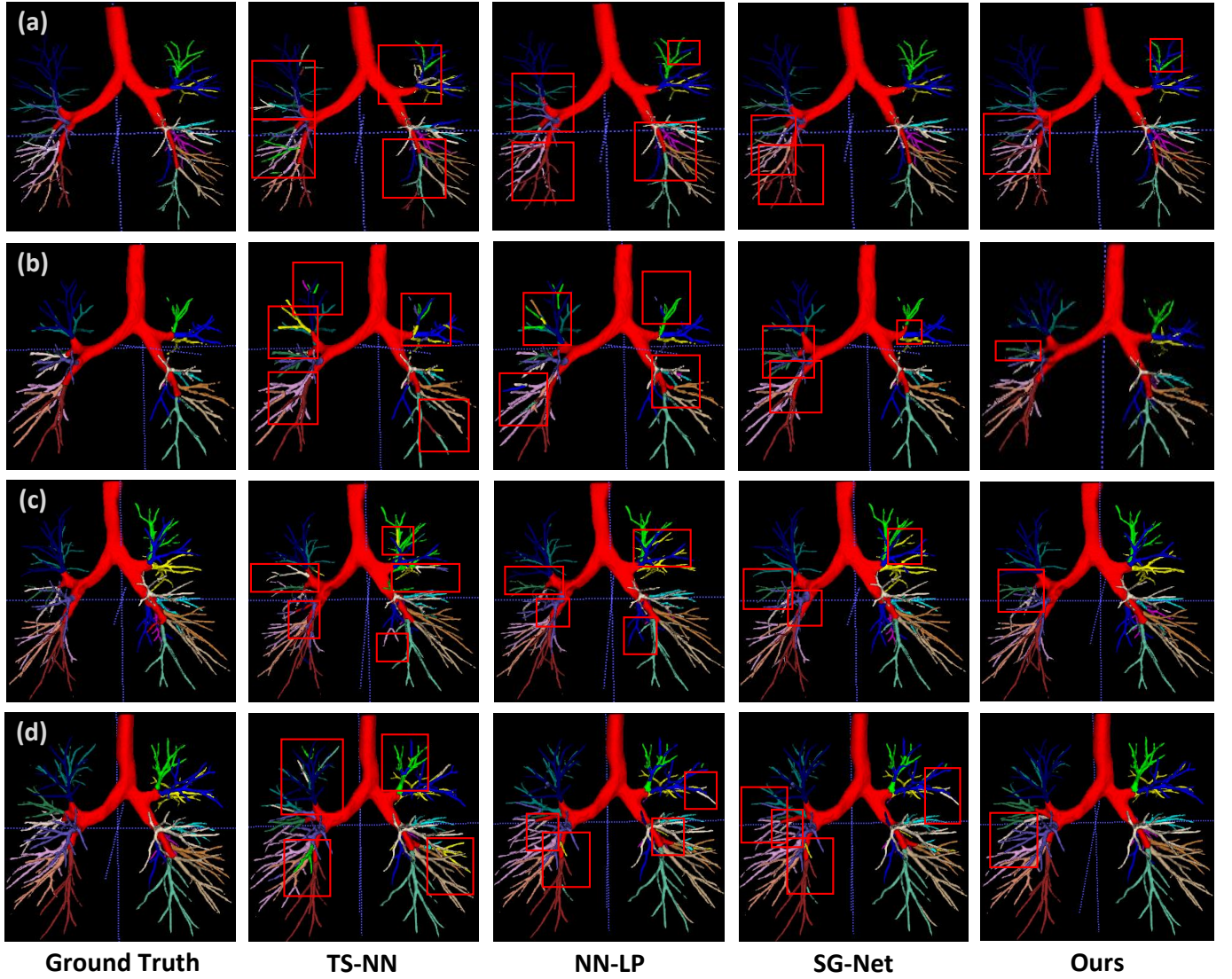


Fig. 5. Qualitative analysis on the BronAtlas benchmark. The misclassified bronchus is bounded by red boxes. TS-NN [7] suffers from errors related to branching variability. LP [31] fails to distinguish the segments with similar angles. SGNet [44] misclassifies the thin bronchus while the proposed BCNet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

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 thus enjoys both the best performance 93.7%, and the same

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±8.5	<0.001
V-GNN			✓		91.4±5.0	0.04
P-GNN		✓			88.2±5.7	<0.001
VN-GNN			✓	✓	91.7±5.5	0.06
PN-GNN		✓		✓	90.7±5.9	0.01
PV-GNN		✓	✓		91.8±6.0	0.07
PVN-GNN		✓	✓	✓	92.4±6.5	0.16
BCNet (GNN)	✓	✓	✓	✓	91.2±6.3	0.03
BCNet (CNN)	✓	✓	✓	✓	93.7±5.5	-

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 (±15.7), precision of 34.3 (±12.2), and particularly low recall of 11.8 (±3.7), resulting in an F1-score of 17.0 (±4.9). 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 (±5.6), precision of 93.4 (±6.1), and recall of 91.7 (±8.2), with an F1-score of 92.5 (±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 preprocessing 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±15.7	34.3±12.2	11.8±3.7	17.0±4.9	<0.001
two-stage	93.6±5.6	93.4±6.1	91.7±8.2	92.5±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.

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.

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±6.1	93.1±5.8	93.7±5.5	93.3±4.9	93.0±6.5
β	0	0.5	1	2	-
F1-score	-	93.5±4.7	93.7±5.5	93.4±5.3	-

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.

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

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,” *American Journal of Respiratory and Critical Care Medicine*, vol. 194, no. 7, pp. 794–806, 2016.
- [2] Y. Qin *et al.*, “Airwaynet: a voxel-connectivity aware approach for accurate airway segmentation using convolutional neural networks,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 212–220, Springer, 2019.
- [3] Y. Kawata *et al.*, “Representation of thoracic n1 lymph nodes group in contrast-enhanced ct images using distance maps based on tracheo-bronchial labeling,” in *Medical Imaging 2023: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 12468, pp. 368–373, SPIE, 2023.
- [4] H. Li, Z. Tang, Y. Nan, and G. Yang, “Human treelike tubular structure segmentation: A comprehensive review and future perspectives,” *Computers in Biology and Medicine*, vol. 151 Pt A, p. 106241, 2022.
- [5] J. Yun *et al.*, “Improvement of fully automated airway segmentation on volumetric computed tomographic images using a 2.5 dimensional convolutional neural net,” *Medical Image Analysis*, vol. 51, pp. 13–20, 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 *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 198–207, Springer, 2017.
- [7] S. A. Nadeem, *A fully automated CT-based airway segmentation and branch labeling algorithm using deep learning and conventional image processing*. PhD thesis, University of Iowa, 2020.
- [8] M. Wang *et al.*, “Automated labeling of the airway tree in terms of lobes based on deep learning of bifurcation point detection,” *Medical & Biological Engineering & Computing*, vol. 58, no. 9, pp. 2009–2024, 2020.
- [9] K. Mori, J.-i. Hasegawa, Y. Suenaga, and J.-i. Toriwaki, “Automated anatomical labeling of the bronchial branch and its application to the virtual bronchoscopy system,” *IEEE Transactions on medical imaging*, vol. 19, no. 2, pp. 103–114, 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?,” *Journal of thoracic disease*, vol. 11, no. 3, p. 972, 2019.
- [11] R. Keuth, M. Heinrich, M. Eichenlaub, and M. Himstedt, “Airway label prediction in video bronchoscopy: capturing temporal dependencies utilizing anatomical knowledge,” *International Journal of Computer Assisted Radiology and Surgery*, pp. 1–9, 2024.
- [12] H. Gong, W. Huang, H. Zhang, Y. Wang, X. Wan, H. Shen, G. Li, and H. Li, “Intensity confusion matters: An intensity-distance guided loss for bronchus segmentation,” in *IEEE International Conference on Multimedia and Expo (ICME)*, 2024.
- [13] H. Gong, G. Chen, R. Wang, X. Xie, M. Mao, Y. Yu, F. Chen, and G. Li, “Multi-task learning for thyroid nodule segmentation with thyroid region prior,” in *2021 IEEE 18th international symposium on biomedical imaging (ISBI)*, pp. 257–261, IEEE, 2021.
- [14] H. Gong, J. Chen, G. Chen, H. Li, G. Li, and F. Chen, “Thyroid region prior guided attention for ultrasound segmentation of thyroid nodules,” *Computers in biology and medicine*, vol. 155, p. 106389, 2023.
- [15] M. Zhang *et al.*, “Multi-site, multi-domain airway tree modeling,” *Medical Image Analysis*, vol. 90, p. 102957, 2023.
- [16] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, Springer, 2015.
- [17] A. Heitz *et al.*, “Lubrav: a new framework for the segmentation of the lung’s tubular structures,” in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pp. 948–952, IEEE, 2021.
- [18] J.-P. Charbonnier, E. M. Van Rikxoort, A. A. Setio, C. M. Schaefer-Prokop, B. van Ginneken, and F. Ciompi, “Improving airway segmentation in computed tomography using leak detection with convolutional networks,” *Medical Image Analysis*, vol. 36, pp. 52–60, 2017.
- [19] F. Isensee, P. F. Jaeger, S. 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, 2021.
- [20] Y. Qin *et al.*, “Learning bronchiole-sensitive airway segmentation cnns by feature recalibration and attention distillation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 221–231, Springer, 2020.
- [21] Y. Qin *et al.*, “Learning tubule-sensitive cnns for pulmonary airway and artery-vein segmentation in CT,” *IEEE Transactions on Medical Imaging*, vol. 40, no. 6, pp. 1603–1617, 2021.
- [22] A. Garcia-Uceda, R. Selvan, Z. Saghir, H. A. Tiddens, and M. de Bruijne, “Automatic airway segmentation from computed tomography using robust and efficient 3-D convolutional neural networks,” *Scientific Reports*, vol. 11, no. 1, pp. 1–15, 2021.
- [23] H. Zheng *et al.*, “Alleviating class-wise gradient imbalance for pulmonary airway segmentation,” *IEEE Transactions on Medical Imaging*, vol. 40, no. 9, pp. 2452–2462, 2021.
- [24] H. Zheng *et al.*, “Refined local-imbalance-based weight for airway segmentation in CT,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 410–419, Springer, 2021.
- [25] Z. Tang, Y. Nan, S. Walsh, and G. Yang, “Adversarial transformer for repairing human airway segmentation,” *IEEE Journal of Biomedical and Health Informatics*, 2023.
- [26] Y. Nan *et al.*, “Fuzzy attention neural network to tackle discontinuity in airway segmentation,” *IEEE Transactions on Neural Networks and Learning Systems*, 2023.
- [27] J. Tschirren, G. McLennan, K. Palágyi, E. A. Hoffman, and M. Sonka, “Matching and anatomical labeling of human airway tree,” *IEEE Transactions on Medical Imaging*, vol. 24, no. 12, pp. 1540–1547, 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 2012: 15th International Conference, Nice, France, October 1–5, 2012, Proceedings, Part III* 15, pp. 147–155, Springer, 2012.
- [29] Y. Liu *et al.*, “Automated anatomical labeling of a topologically variant abdominal arterial system via probabilistic hypergraph matching,” *Medical Image Analysis*, vol. 75, p. 102249, 2022.
- [30] P. Lo, E. M. van Rikxoort, J. Goldin, F. Abtin, M. de Bruijne, and M. Brown, “A bottom-up approach for labeling of human airway trees,” *MICCAI Int. WS. Pulm. Im. Anal*, 2011.
- [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 *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 230–239, Springer, 2019.

- [32] W. Yu *et al.*, “Tnn: Tree neural network for airway anatomical labeling,” *IEEE Transactions on Medical Imaging*, vol. 42, no. 1, pp. 103–118, 2022.
- [33] M. Defferrard, X. Bresson, and P. Vandergheynst, “Convolutional neural networks on graphs with fast localized spectral filtering,” in *NeurIPS*, 2016.
- [34] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, “Graph attention networks,” in *International Conference on Learning Representations*, 2018.
- [35] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, “How powerful are graph neural networks?,” in *ICLR*, 2019.
- [36] X. Song, J. Li, and X. Qian, “Diagnosis of glioblastoma multiforme progression via interpretable structure-constrained graph neural networks,” *IEEE Transactions on Medical Imaging*, vol. 42, no. 2, pp. 380–390, 2022.
- [37] Y. Li, Y. Shen, J. Zhang, S. Song, Z. Li, J. Ke, and D. Shen, “A hierarchical graph v-net with semi-supervised pre-training for histological image based breast cancer classification,” *IEEE Transactions on Medical Imaging*, 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 bioinformatics*, vol. 23, no. 1, p. 394, 2022.
- [39] H. Gong, Y. Zhang, C. Dong, Y. Wang, G. Chen, B. Liang, H. Li, L. Liu, J. Xu, and G. Li, “Unbiased curriculum learning enhanced global-local graph neural network for protein thermodynamic stability prediction,” *Bioinformatics*, vol. 39, no. 10, p. btad589, 2023.
- [40] G. Corso, H. Stark, S. Jegelka, T. Jaakkola, and R. Barzilay, “Graph neural networks,” *Nature Reviews Methods Primers*, vol. 4, no. 1, p. 17, 2024.
- [41] A. Garcia-Uceda Juarez, R. Selvan, Z. Saghir, and M. d. Bruijne, “A joint 3D unet-graph neural network-based method for airway segmentation from chest CTs,” in *International Workshop on Machine Learning in Medical Imaging*, pp. 583–591, Springer, 2019.
- [42] R. Selvan *et al.*, “Graph refinement based airway extraction using mean-field networks and graph neural networks,” *Medical Image Analysis*, vol. 64, p. 101751, 2020.
- [43] T. Zhao and Z. Yin, “Airway anomaly detection by prototype-based graph neural network,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 195–204, Springer, 2021.
- [44] Z. Tan, J. Feng, and J. Zhou, “SGNet: Structure-aware graph-based network for airway semantic segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 153–163, Springer, 2021.
- [45] W. Lou, X. Wan, G. Li, X. Lou, C. Li, F. Gao, and H. Li, “Structure embedded nucleus classification for histopathology images,” *IEEE Transactions on Medical Imaging*, 2024.
- [46] W. Lou, G. Li, X. Wan, and H. Li, “Cell graph transformer for nuclei classification,” in *Thirty-Eighth AAAI Conference on Artificial Intelligence*, AAAI 2024, February 20-27, 2024, Vancouver, Canada, pp. 3873–3881, AAAI Press, 2024.
- [47] P. Lo *et al.*, “Extraction of airways from ct (exact’09),” *IEEE Transactions on Medical Imaging*, vol. 31, no. 11, pp. 2093–2107, 2012.
- [48] S. G. Armato III *et al.*, “The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans,” *Medical Physics*, vol. 38, no. 2, pp. 915–931, 2011.
- [49] J. R. Williams, “The declaration of helsinki and public health,” *Bulletin of the World Health Organization*, vol. 86, pp. 650–652, 2008.
- [50] T.-C. Lee, R. L. Kashyap, and C.-N. Chu, “Building skeleton models via 3-D medial surface axis thinning algorithms,” *CVGIP: Graphical Models and Image Processing*, vol. 56, no. 6, pp. 462–478, 1994.
- [51] S. A. Nadeem, E. A. Hoffman, J. C. Sieren, A. P. Comellas, S. P. Bhatt, I. Z. Barjaktarevic, F. Abtin, and P. K. Saha, “A CT-based automated algorithm for airway segmentation using freeze-and-grow propagation and deep learning,” *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp. 405–418, 2020.
- [52] W. L. Hamilton, Z. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” in *Advances in neural information processing systems*, vol. 30, 2017.
- [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 *ICML*, 2021.
- [54] Y. Rong, W. Bing Huang, T. Xu, and J. Huang, “DropEdge: Towards deep graph convolutional networks on node classification,” in *ICLR*, 2020.