Modified Attention Unet for Lung Airway Accurate Segmentation with Warmup Cosine Annealing Learning Rate

Boyu Zhang, Yinglan Kuang, Huajia Wang, Xi Liu, Xing Lu

A.I. R&D Center, Sanmed Biotech Inc. Zhuhai, Guangdong, China

lv.xing@sanmedbio.com

Abstract

To accurately segment lung airway from low dose CT scans, a modified attention U-net is proposed for the Airway Tree Modeling Challenge 2022 and trained with warmup cosine annealing learning rate strategy on a 299 dataset provided by the organizer. The dice score of our model achieved 0.946 on the validation dataset.

Keywords

Airway segmentaion, Modified Unet, Coordinate Attention, Project & Excite module

Introduction

Accurate lung airway segmentation is of great importance for the diagnosis and analysis of pulmonary diseases, which can provide effective anatomic information for pulmonary surgery[1,2].

However, Airway segmentation is a critical step in the analysis of lung diseases, including asthma, bronchiectasis, and emphysema. Accurate segmentation based on X-ray computed tomography (CT) can quantitatively measure airway size and wall thickness to reveal abnormalities in patients with chronic obstructive pulmonary disease (COPD). In addition, the extraction of patient-specific airway models from CT images is necessary for the navigation of bronchoscopy assisted surgery. However, manual notation is time-consuming, error-prone, and highly dependent on clinician expertise due to the fine-grain structure of the lung airways.

For the Airway Tree Modeling Challenge 2022(ATM22) competition of the MICCAI 2022, 500 CT scans were collected from multi-sites[3-5]. The airway tree structures are carefully labeled by three radiologists with more than five years of professional experience. The intra-class imbalance among the trachea, main bronchi, lobar bronchi, and distal segmental bronchi affects the segmentation performance of peripheral bronchi. In conclusion, we encourage the participating teams to design robust algorithms, which can extract the airway tree structure with high topological completeness and accuracy for clinical use.

Method

Data

For the Validation Phase 1, 300 CT scans were provided as public training dataset. According to the reminder from the organizer, one abnormal scan was removed, with 299 scans as training phase dataset. In our training process, 270 scans (from ATM_001_0000 to ATM_642_0000) were used for model training and the rest 29 scans were used for validation to evaluate the model performance.

Preprocess

In preprocessing period, the window width and level of all scans is set as 1600 and -600, retrospectively. The entry of the thoracic cavity into the model is likely to cause wrong prediction, so the image processing method is used to extract the mask of the left and right lobes. Since the annotated length of the main airway of the contest sample data is not uniform, the main airway part is retained when taking the lung mask. According to the box generated by lung mask, the lung data in CT images were intercepted. The CT values were normalized to the gray scale range (0-255). Due to limited training resources, the preprocessed CT images were cut into 128*128*128 small size images and then input into the network.

Model architecture

The airway segmentation network we used for this competition is modified according to the classical U-net. As demonstrated in Fig 1.

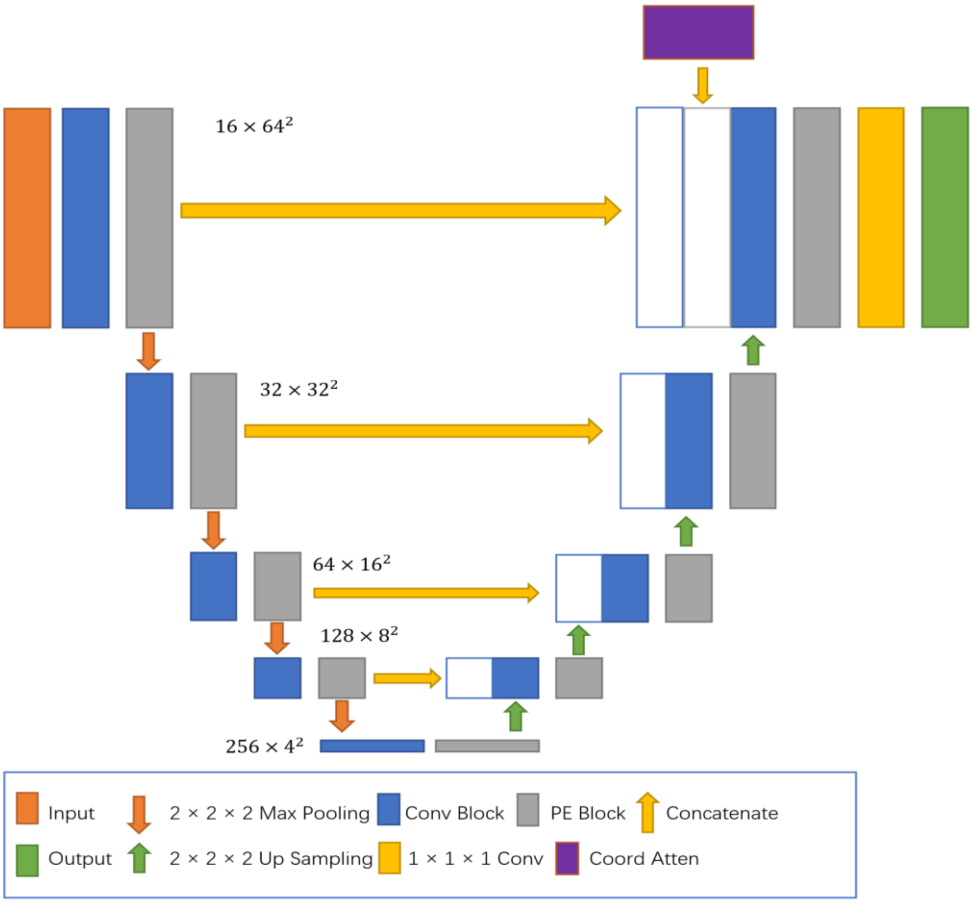

Fig 1. Network architecture for airway segmentation

In our network, Project & Excite (PE) module is added behind each convolution layer. The module first projected feature maps along each axis to get three feature maps with shapes of D×1×1×C, 1×H×1×C and 1×1×W×C, then expanded them and added them together, and finally got an attention map with shape of D×H×W×C. The PE block is more relevant to channel attention and assigns a voxel-level attention coefficient to each channel. It helps the network to learn the important feature information and improve the generalization ability of the model.

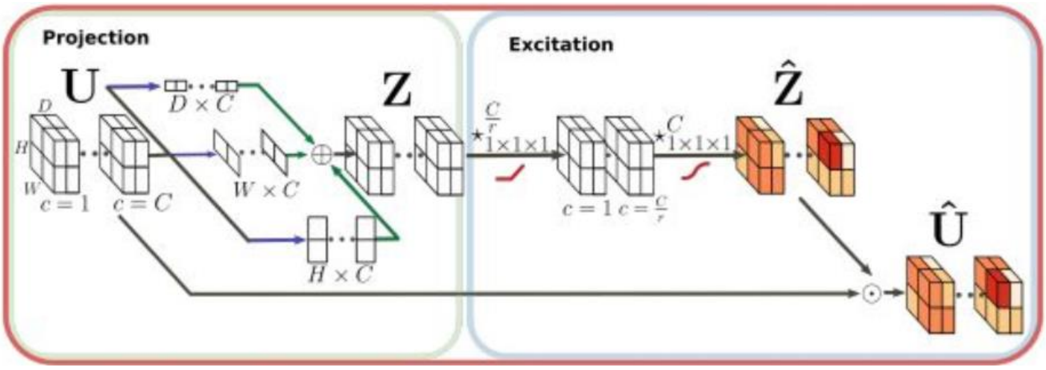

Fig 2. PE module used in this study.

And then at the end of the network we put the Coordinate Attention mechanism. The network has two inputs, patch and Coord. Coord retains the location information of its corresponding patch in the whole image. According to the relative position of each pixel in the patch in the original image, normalization is made in the range of [-1,1] in the three axes. Coord dimensions for [coordinate, the depth, height and width), including the dimensions represent the spatial dimension, coordinate information matrix of the layer, the height of the coordinate matrix, the width of the coordinate matrix. The input dimension used in this method is [3,128,128].

The final loss function is composed of Dice loss and Focal loss.

$$L_{seg} = L_{Dice} + L_{Focal} \quad [1]$$

Model Training and Inference

Model was implemented with torch 1.4 and trained on a workstation with two Tesla T4 GPU(32 GB in total) . CosineAnnealingWarmupRestarts module was used as training learning rate, with the highest LR of 0.01 and lowest boundary of 1e-5. The cycle period is of 20 epochs and decay ratio for each cycle is 0.5. Training and evaluation for each epoch was recorded with tensorboardX module and is demonstrated in Fig 3. Models with best performance on the evaluation set and every 10 echoes were saved for further inference.

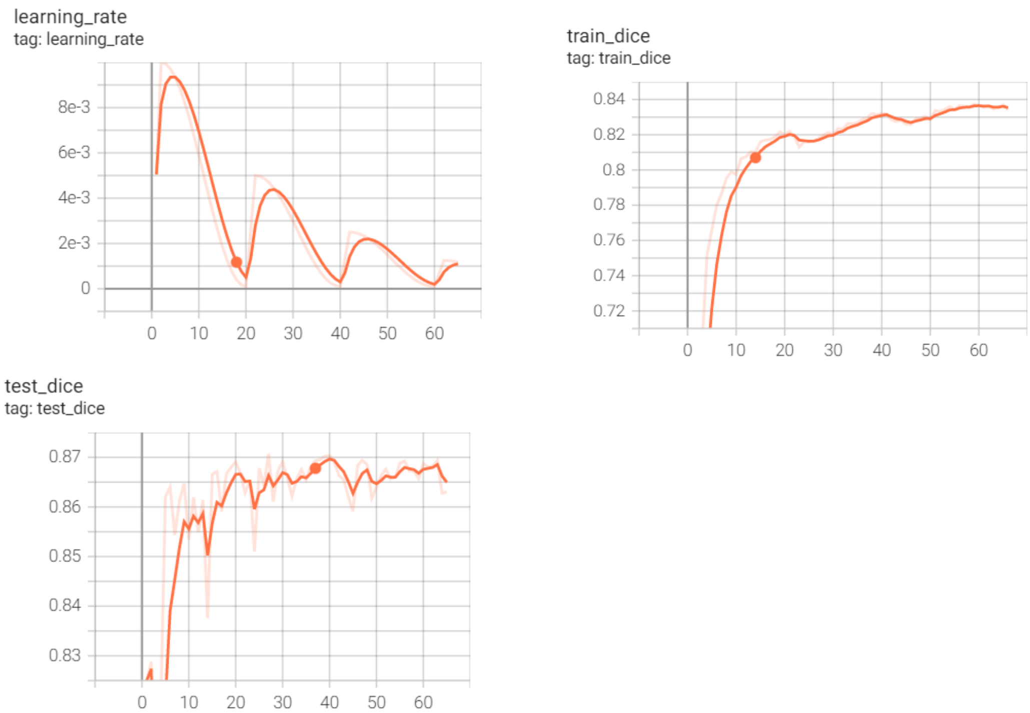

Fig 3. Training recorded for our model.

During model inference on the validation dataset provided by the organizer, the connected component labeling algorithm(CCLA) was implemented to remove some isolated predicted area, which mostly are false positive detection of the airway. By using the measure module of the skimge, with label() function to obtain every connected components. By remove the background connected component and tiny connected components, only the connected component with maximum volume remained as the predicted results.

Results

Finally, we selected models that saved at epoch 20th and 71st to submit to the lead board. As shown in table 1, the model at epoch 20th and 71st obtained dice score of around 0.87 for the model performance and improved to above 0.936 after CCLA at our own validation dataset, but get higher validation score of 0.9463 and 0.9461, retrospectively, on the public validation dataset.

Table 1. Submission results

Submit	Model	Evaluation Dice	After post-process	Final validation
1st	Epoch 40	0.9335	0.9366	0.9463
2ed	Epoch 71	0.9343	0.9375	0.9461

Conclusion

In this study, a modified 3D Unet with PE module and coordinate attention model was proposed and used for the airway segmentation. By warmup cosine annealing learning rate training strategy, the dice score is evaluated as 0.946 on the validation phase 1 of the ATM22 competition.

Reference

[1] Zhang M, Yu X, Zhang H, et al. FDA: Feature Decomposition and Aggregation for Robust Airway Segmentation[M]//Domain Adaptation and Representation Transfer, and Affordable Healthcare and AI for Resource Diverse Global Health. Springer, Cham, 2021: 25-34.

[2] Zheng H, Qin Y, Gu Y, et al. Alleviating class-wise gradient imbalance for pulmonary airway segmentation[J]. IEEE Transactions on Medical Imaging, 2021, 40(9): 2452-2462.

[3] Yu W, Zheng H, Zhang M, et al. BREAK: Bronchi Reconstruction by gEodesic transformation And sKeleton embedding[C]//2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). IEEE, 2022: 1-5.

[4] Qin Y, Chen M, Zheng H, et al. Airwaynet: a voxel-connectivity aware approach for accurate airway segmentation using convolutional neural networks[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2019: 212-220.

[5] Y. Qin, H. Zheng, Y. Gu, X. Huang, J. Yang, L. Wang, F. Yao, Y.-M.Zhu, G.-Z. Yang, Learning

Tubule-Sensitive CNNs for PulmonaryAirway and Artery-Vein Segmentation in CT, IEEE Transactions onMedical Imaging 40 (2021) 1603–1617.