Segmentation of Pulmonary Arteries From Volumetric Chest CT Scans with Deep Learning Methods

Haichao Peng 23020221154172 Informatics class, Jinyu He 23020221154160 Informatics class, Zhibin Huang 23020221154163 Informatics class, Yue Ma 23020221154170 Informatics class, Gaiqun Zhang 33320221150333 AI class

Abstract

Extraction, segmentation and analysis of pulmonary arteries from computed tomography (CT) images of the human chest is an important topic for a wide range of applications in medical image analysis. Traditionally, the radiologist's subjective diagnosis is largely dependent on his or her experience. Besides, some methods have been proposed to automatic segment the lungs' arteries, like region growing and other traditional machine learning methods, but with relatively low accuracy. In this paper, we employ deep learning methods to improve segmentation results. We evaluate these methods on 50 CT images, with the experimental results showing that deep learning methods gain a better segmentation results and a faster speed as well. Moreover, we propose a noval refinement network architature based on U-Net to further improve segmentation preformance.

Introduction

Pulmonary arteries segmentation is a topic of high interest in the field of medical image analysis: accurate vascular analysis has extremely important research and application value for treatment planning and clinical effect evaluation. Pulmonary arteries segmentation is a basis for common pulmonary vascular diseases diagnosis(Van Rikxoort et al. 2009) such as lobectomy and pulmonary embolism, which is a more complex and difficult task as 18 categories should be classified, as shown in Figure 1.

Lung CT is mainly for the lung region and mediastinal region image sequence slice. In the observation of pulmonary blood vessels, the image data of the mediastinal window of the lung has the most research value, so lung CT is now one of the most important means to evaluate the pulmonary function. At present, in clinical diagnosis and treatment, professional doctors mainly diagnose pulmonary hypertension by observing the cross-sectional area of grade 4 to 5 blood vessels in the original CT images and observing whether the blood vessels are blocked, which requires the pulmonary blood vessels to be marked first. Currently, the labeling of pulmonary vessels usually relies heavily on manual operation by professional physicians. Due to the amount of data generated by modern Computed tomography angiography (CTA), manual labeling is time-consuming, labor-intensive and error-prone. In addition, professional doctors tend to label some iconic and representative pictures in order to save time, which is easy to ignore a lot of insignificant but important diagnostic information, resulting in blurred details of lesions, and even lead to misdiagnosis.

The segmentation of pulmonary arteries in CT images is the basis and key of computer-aided detection and diagnosis of pulmonary diseases such as pneumonia and pulmonary embolism(McAdams, Erasmus, and Winter 1995).

Lung segments are divided according to the branches of the bronchus. There are fibrous septa between each lung segment, which can be used to diagnose some diseases. However, if the trachea extends to the periphery, it is difficult to identify it on CT. Therefore, we divided the arteries into 18 categories and further divided the lungs into 18 segments, which is helpful for clinical segmental pulmonary resection.

However, the human lung, on the other hand, is an exchange site for gases produced by metabolism and is rich in trachea and vascular tissue, so its structure is relatively complex. At the same time, due to the influence of noise, volume effect and other factors, CT images may appear poor contrast, blurred boundary and other problems. Moreover, the pulmonary arteries and veins intertwine with each other, which further increases the difficulty of segmentation.

Therefore, it is of great significance to provide more efficient and accurate medical diagnosis if we can automate and accurately label pulmonary arteries with the help of efficient algorithms and auxiliary tools.

Related Work

Pulmonary segments segmentation has always been a hot research problem, while pulmonay arteries segmentation can be the first step of it. Pulmonary arteries segmentation based on CT images also can help doctors diagnose many pulmonary diseases more accurately and effectively, such as pulmonary embolism and primary pulmonary hypertension. However, CT images have noise, small size of pulmonary vessels and blurred edges, which make segmentation of pulmonary arteries a challenging problem. Besides, the study of pulmonary arteries segmentation is relatively few compared with other studies, such as pulmonary vessels segmentation, which is an easier task with two categories(vein and artery). In recent years, researchers have tried a variety of methods to accurately segment pulmonary blood vessels, such as threshold segmentation, region growing method, Hessian filter, graph cut algorithm(Zhai, Staring, and Stoel 2016), level

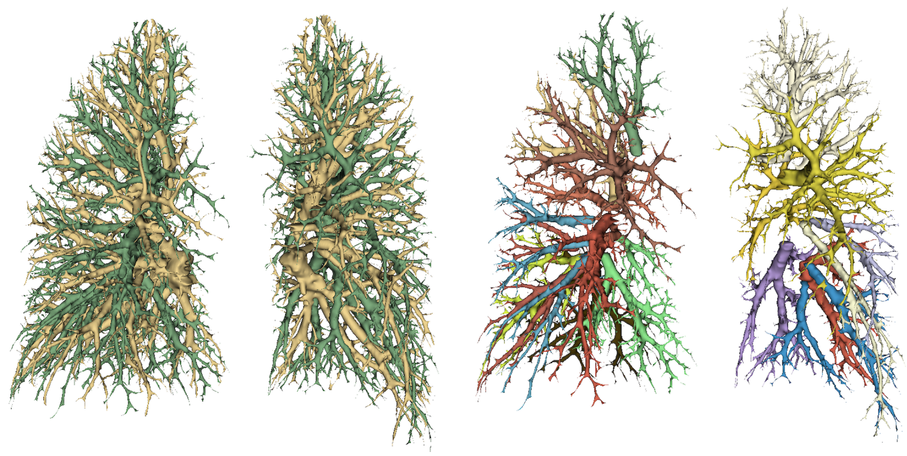

Figure 1: The left is a schematic of pulmonary vessel segmentation (2 categories), while a more complex and difficult task, pulmonary arteries segmentation (18 categories), is on the right.

set based segmentation method, and machine learning methods:

The graph cutting method

The graph cutting method regards the segmentation of blood vessels as a marking problem. According to the node adjacency and its weight, the voxel nodes are marked as targets or backgrounds to achieve segmentation. Chen et al. proposed a liver vessel segmentation method based on region graph cut, which was initialized by fast shift clustering. Freiman et al. proposed a graph-cut method for carotid artery segmentation by combining vessel distribution and intensity into the cost function. The scan image was segmented into several blocks with small overlap, and the graph cut was calculated for each block, and these classification results were combined. However, the graph cut method can only achieve local optimum, which is good for the image with high contrast, but not satisfactory for the image with low contrast.

Supervised machine learning methods

In addition to the classical methods above, supervised machine learning methods can also be used to segment pulmonary blood vessels. Kiros et al. used logistic regression classifiers with K-means features, which were obtained by Orthogonal Matching Pursuit (OMP)(Kiros et al. 2014), to segment pulmonary vessels. OMP algorithm will iterate K cycles, but in general, K is often unknown, which may lead to inconvenience in the training process. BowenZhao et al. used a combination of sparse autoencoder and random forest to segment pulmonary vessels, where the autoencoder can learn the structure of pulmonary vessels by imposing sparsity on hidden units during training.

Deep learning method

CNN-based Segmentation Networks : Since the introduction of the seminal U-Net(Ronneberger, Fischer, and Brox 2015), CNN-based networks have achieved state-of-the-art results on various 2D and 3D various medical image segmentation tasks. For volume-wise segmentation, tri-planar architectures are sometimes used to combine three-view slices for each voxel, also known for 2.5D methods(Roth et al. 2014). In contrast, 3D approaches directly utilize the full volumetric image represented by a sequence of 2D slices or modalities. The intuition of employing varying sizes was followed by multi-scan, multi-path models to capture downsampled features of the image. In addition, to exploit 3D context and to cope with limitation of computational resource, researchers investigated hierarchical frameworks. Some efforts proposed to extract features at multiple scales or assembled frameworks. Roth et al. proposed a multi-scale framework to obtain varying resolution information in pancreas segmentation. These methods provide pioneer studies of 3D medical image segmentation at multiple levels, which reduces problems in spatial context and low-resolution condition. Despite their success, a limitation of these networks is their poor performance in learning global context and long-range spatial dependencies, which can severely impact the segmentation performance for challenging tasks.

Method

In this section, we will first briefly review the process of pulmonary arteries segmentation based on deep learning and the basic UNet model(Ronneberger, Fischer, and Brox 2015), then we introduce our improved version of UNet, called Refinement UNet, as shown in Figure 2.

Pulmonary Arteries Segmentation

Since the human lung can be divided into 18 segments, the entire pulmonary artery can be divided into 18 classes of arteries. Based on this, for any given lung CT, we need to segment the arteries and classify them into 18 classes at the pixel level, plus the non-artery part, for a total of 19 classes. More specifically, given a single-channel CT with input size $H \times W \times D$ (denote height, width and depth respectively), i.e. $x \in R^{1\times H\times W\times D}$, it is input into a deep neural

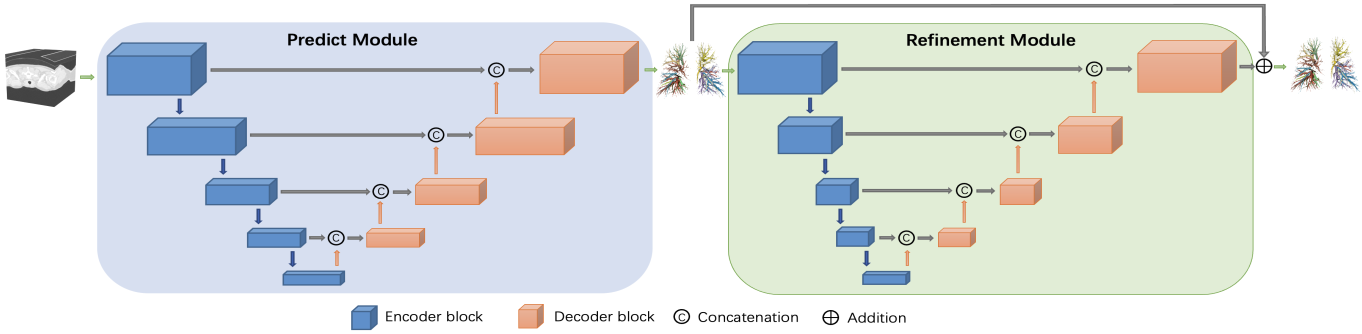

Figure 2: Architecture of our proposed Refinement UNet.

network (such as 3DUNet(Ronneberger, Fischer, and Brox 2015), 3DU2Net(Qin et al. 2020), UNETR(Hatamizadeh et al. 2022b), etc.), and the final output is a predicted result with shape (19, H, W, D), i.e. $\hat{y} \in R^{19 \times H \times W \times D}$.

UNet

UNet(Ronneberger, Fischer, and Brox 2015) is a deep learning model that is designed for image segmentation and other morphological tasks. It consists of two parts in its network architecture: a long downsampling path and a short upsampling path. The downsampling path, also known as the "encoder," reduces the size of the image using max pooling layers and regular convolutional layers, and in the process, extracts features. The upsampling path, also known as the "decoder," enlarges the image using transposed convolutional layers, also known as "deconvolution," and uses the feature maps to restore the details of the image.

UNet(Ronneberger, Fischer, and Brox 2015) also uses a technique called "skip connections," which allows the model to skip over multiple layers between the downsampling path and the upsampling path, allowing for better information flow. In this way, each layer in the upsampling path can simultaneously access features from lower layers and higher layers.

However, UNet is mostly used in 2D inputs while our inputs are CT images which are 3D volumetric data. Therefore, we need to convert UNet, a 2D network by default, to 3DUNet. Then, 3D CT image data can be allowed to feed into it.

Refinement UNet

In our task, it is easy to predict the stems of pulmonary arteries, but difficult for the periphery vessels. Inspired from (Islam et al. 2017; Deng et al. 2018), we propose a refinement module which is designed as a residual block to refines the predicted coarse prediction probability maps S_{coarse} by learning the residuals $S_{residual}$, it can be fomulated as

$$S_{refined} = S_{coarse} + S_{residual}$$

In our Refinement UNet, we use a small 3DUNet with less channels in each stage to predict the coarse prediction maps, inspired by U2Net which achieves better performance with lower complexity in our task. Unlike (Islam et al. 2017; Deng et al. 2018), we use an another relative smaller 3DUNet as our refinement module. We mainly consider two aspects: 1) shallow refinement module has a small receptive field, like (Peng et al. 2017); 2) multi-scale contexts may help to capture high level information for refinement. Additionally, the input of refinement module is combination of original input and the coarse predicition maps. Finally, refined prediction maps come out by adding the residual outputs and the coarse prediction results.

Loss Function

Our loss function is based on the commonly used soft dice loss(Milletari, Navab, and Ahmadi 2016). It is defined as:

$$\mathcal{L}(y, \hat{y}) = 1 - \frac{1}{I} \sum_{i}^{I} \frac{2 \sum_{v=1}^{V} y_{v,i} \cdot \hat{y}_{v,i}}{\sum_{v=1}^{V} y_{v,i}^2 + \sum_{v=1}^{V} \hat{y}_{v,i}^2}$$

where, I denotes the number of classes; V denotes the number of voxels; $y_{v,i}$ and $\hat{y}_{v,i}$ denote the ground truths and predicted probabilities at voxel v for class i, respectively.

Experiments

Dataset

As few people who work on segmentation of the pulmonary arteries based on lung segmentation, there is no relevant public dataset available. Thus, we selected a portion of the lung CT data provided by LUNA-16 and labeled it with 18 classes of arteries, resulting in a total of 50 usable sets of data. Each CT set consists of between 80 and 320 slices, each with a size of 512×512, i.e. $X : x \in R^{512 \times 512 \times D}, D \in [80, 320]$. In data augmentation phase, we pre-process each CT scan independently by normalizing the intensities in the range of [-1200, 400] HU to [0, 1] with patch cropping of $192 \times 160 \times 160$ and spacing of [1.0, 1.0, 1.6]. Besides, we augment data by using random rotating, flipping.

It is worth noting that, due to the scarcity of medical imaging data, we actually do not need as large a dataset as other fields for this pixel-level dense segmentation task, as the data patterns are relatively simple. In recent years, the datasets used in segmentation-related works on CT and MRI (such as UNETR and nnFormer) are very small. Therefore, 50 sets of data are sufficient.

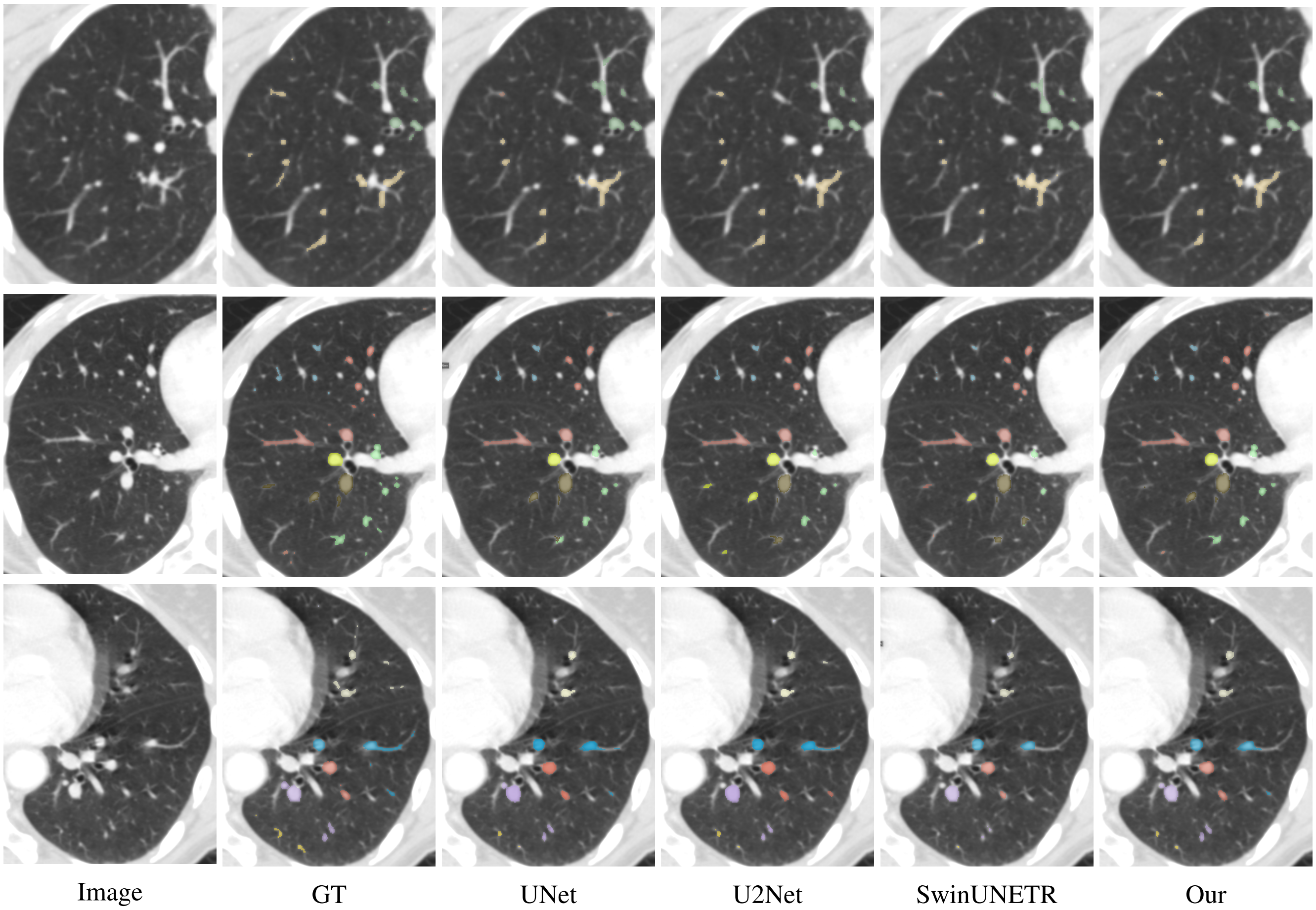

Figure 3: Qualitative comparison on pulmonary arteries segmentation. Here, we compare our Refinement UNet with existing methods: UNet(Ronneberger, Fischer, and Brox 2015), U2Net(Qin et al. 2020), SwinUNETR(Hatamizadeh et al. 2022a). The different types of arteries are distinguished by different colors. Our Refinement UNet achieves promising segmentation performance by accurately segmenting the arteies. Best viewed zoomed in.

In this experiment, we divided the dataset into a training set (40 subjects) and a test set (10 subjects). The model is trained on the training set and the model with the highest accuracy on the test set is saved as our final model.

Implementation Details

We implement our approach in Pytorch v1.12.1 and using the MONAI libraries. The models are trained using a single NVIDIA GeForce RTX 3090Ti 24GB GPU with input 3D pathes of size $192 \times 160 \times 160$ for 60000 iterations. Because of the limited memory, our batch size is only 4. Specifically, each epoch equals 1-10 iterations, so we train each model for 6000 epochs. Besides, we use Adam optimizer with learning rate of $1e^{-3}$ and weight decay of $1e{-}3$, and cosine annealing restart learning rate schedule to dynamically adjust learning rate. During the testing phase, we adopt slide windows method with size equals that in traning phase to predict the whole segmentation for too large size of CT images and limited GPU memory.

Evaluations Metric

We measure the performance of models based on Dice SImilarity Score (DSC). DSC measures the overlap between the volumetric segmentation predictions and voxels of ground truths, it is defined as follows:

$$DSC(Y,P) = \frac{2|Y \cap P|}{|Y| + |P|} = \frac{2\,Y \cdot P}{Y^2 + P^2}$$

where, Y and P denote the ground truths and output probabilities for all voxels, repectively.

Quantitative Evaluations

Table 1 shows the results on the pulmonary arteries segmentation task. We report the segmentation performance using DSC metric. In addition, we report the model complexity in terms of parameters and FLOPs for each method. The segmentation performance is reported with single model accuracy and without utilizing any pre-training, model ensemble or additional data. The original UNet(Ronneberger, Fischer, and Brox 2015) approach achieves a DSC of 72.93%.

Other popular methods, U2Net(Qin et al. 2020) and SwinUNETR(Hatamizadeh et al. 2022a) achieve DSC of 73.74% and 69.58%. respectively. On this task, our Refinement UNet outperforms other method above by achieving a DSC of 76.16%. Further, Refinement UNet achieves this improvement in segmentation performance by significantly reducing the model complexity compared to the original UNet and transformer based models.

In Figure 3, we visualize some segmentation results of our Refinement UNet, UNet(Ronneberger, Fischer, and Brox 2015), U2Net(Qin et al. 2020) and SwinUNETR(Hatamizadeh et al. 2022a) on pulmonary ariteries segmentation task. Compared to other methods, our method can greatly reduce the number of false positive predictions. One typical example is the second raw. We can see that other methods produce a large number of wrong pixels to the green class which actually is supposed to be classified to vein or background. These results verfity that our Refinement UNet has great potential in this task.

Methods	Average	Left Lung	Right Lung	Params	Flops
UNet	72.93	72.85	73.03	19.26M	427.48B
U2Net	73.74	73.54	73.99	3.56M	2495.57B
SwinUNETR	69.58	70.31	68.67	15.11M	460.98B
Re-UNet	**76.16**	**76.74**	**75.44**	3.99M	744.49B

Table 1: Comparison with deep learning models on pulmonary arteries segmentation. The evaluation metric is DSC in (%). Best results are bolded. Our proposed Refinement UNet achieves favorable segmentation performance against existing methods, while being considerably reducing the model complexity.

Conclusion

In this paper, we conduct pulmonary segments based arteries segmention on 50 self-labeled CT images by employing deep learning methods, like 3DUNet(Ronneberger, Fischer, and Brox 2015), 3DU2Net(Qin et al. 2020) and SwinUNETR(Hatamizadeh et al. 2022a). Furthermore, we propose a refinement UNet, which is a predict-refine rachitecture, which consists of two components: a prediction network and a refinement module. Experimental results demonstrate that our proposed model gains better results. It is noted that the refinement module can be easily extended or adapted to the task by replacing either the predicting network or the refinement module.

References

Deng, Z.; Hu, X.; Zhu, L.; Xu, X.; Qin, J.; Han, G.; and Heng, P.-A. 2018. R3net: Recurrent residual refinement network for saliency detection. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, 684–690. AAAI Press Menlo Park, CA, USA.

Hatamizadeh, A.; Nath, V.; Tang, Y.; Yang, D.; Roth, H. R.; and Xu, D. 2022a. Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In *International MICCAI Brainlesion Workshop*, 272–284. Springer.

Hatamizadeh, A.; Tang, Y.; Nath, V.; Yang, D.; Myronenko, A.; Landman, B.; Roth, H. R.; and Xu, D. 2022b. Unetr: Transformers for 3d medical image segmentation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 574–584.

Islam, M. A.; Kalash, M.; Rochan, M.; Bruce, N. D.; and Wang, Y. 2017. Salient Object Detection using a Context-Aware Refinement Network. In *BMVC*.

Kiros, R.; Popuri, K.; Cobzas, D.; and Jägersand, M. 2014. Stacked Multiscale Feature Learning for Domain Independent Medical Image Segmentation. In *MLMI*.

McAdams, H. P.; Erasmus, J.; and Winter, J. A. 1995. Radiologic manifestations of pulmonary tuberculosis. *Radiologic Clinics of North America*, 33(4): 655–678.

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

Peng, C.; Zhang, X.; Yu, G.; Luo, G.; and Sun, J. 2017. Large kernel matters–improve semantic segmentation by global convolutional network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4353–4361.

Qin, X.; Zhang, Z.; Huang, C.; Dehghan, M.; Zaiane, O. R.; and Jagersand, M. 2020. U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern recognition*, 106: 107404.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234–241. Springer.

Roth, H. R.; Lu, L.; Seff, A.; Cherry, K. M.; Hoffman, J.; Wang, S.; Liu, J.; Turkbey, E.; and Summers, R. M. 2014. A new 2.5 D representation for lymph node detection using random sets of deep convolutional neural network observations. In *International conference on medical image computing and computer-assisted intervention*, 520–527. Springer.

Van Rikxoort, E. M.; De Hoop, B.; Van De Vorst, S.; Prokop, M.; and Van Ginneken, B. 2009. Automatic segmentation of pulmonary segments from volumetric chest CT scans. *IEEE transactions on medical imaging*, 28(4): 621–630.

Zhai, Z.; Staring, M.; and Stoel, B. C. 2016. Lung vessel segmentation in CT images using graph-cuts. In Styner, M. A.; and Angelini, E. D., eds., *Medical Imaging 2016: Image Processing*, volume 9784 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 97842K.